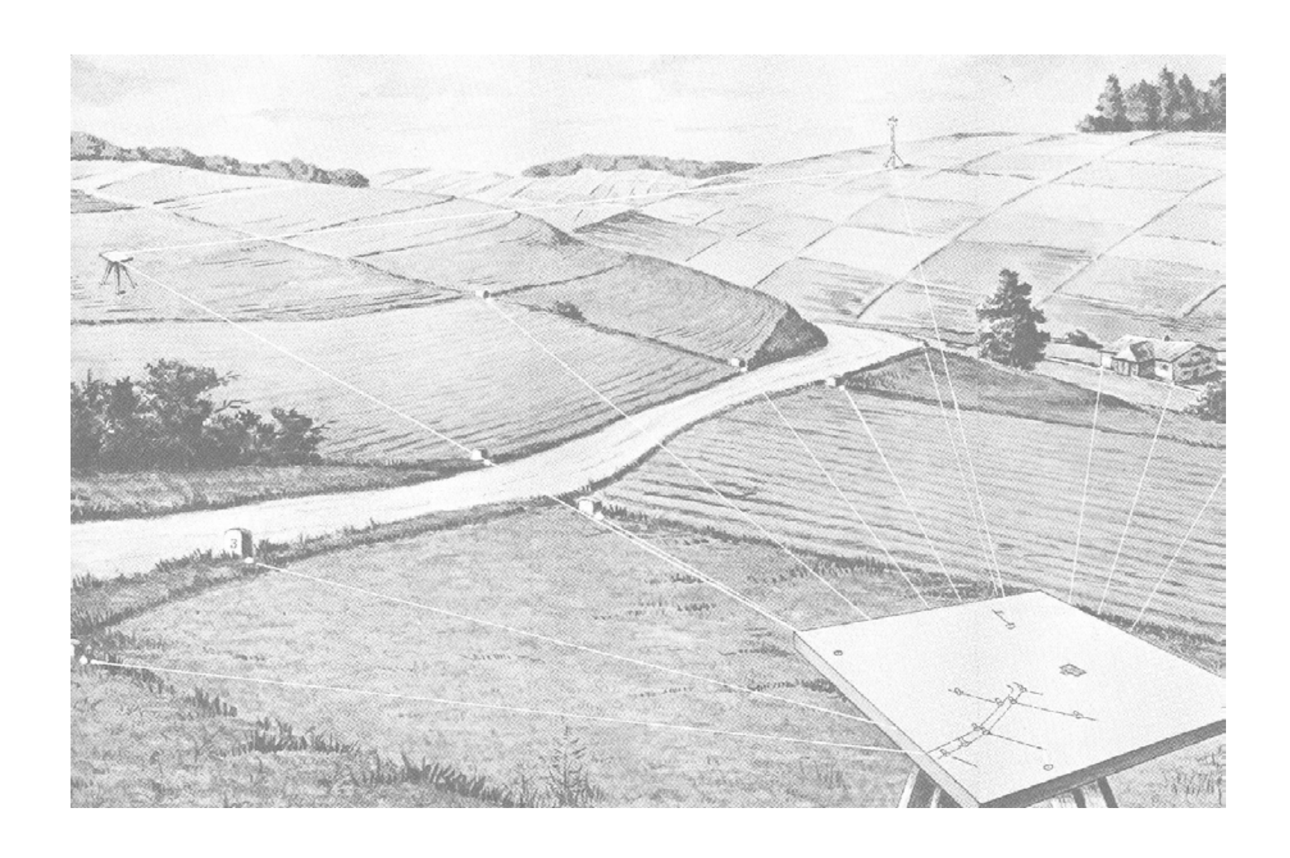


Hochschule München

Fakultät für Geoinformation



Geoinformatik Teil 1

Geoinformatik

Bachelorstudiengang

Geoinformatik und Satellitenpositionierung

Prof. Dr. G. Lother 2012

V12.2

Modul G1: Geoinformatik (GI)

Modul G2: Geoinformationssysteme (GIS)

Modul G3: Geodateninfrastruktur (GDI)

Schwerpunkte:	**GDI-Komponenten (GDI = GDB + GWS)**
	Qualitätsmerkmale von Geodaten

GI - Modul Geoinformatik (GI) 4 SWS [Vorlesung 2SWS + Übungen 2SWS]: „Die Geoinformatik setzt sich mit dem Wesen der Geoinformation, mit ihrer Bereitstellung in Form von Geodaten und mit den darauf aufbauenden Anwendungen auseinander. Die daraus gewonnenen Erkenntnisse münden in die Technologie der Geoinformationssysteme (GIS)."

Den Schwerpunkt der Vorlesung Geoinformatik bilden Datenstrukturen und Konzepte, die für die Abbildung und Organisation von Geoinformation im Rechner benötigt werden. In den Kapiteln Informatikgrundlagen, Datenerfassung und kartographische Visualisierung wird lediglich ein Bezug zu den jeweiligen Fachvorlesungen hergestellt.

Literaturauswahl

Bartelme, N. (Graz 1989) **GIS Technologie** – Geoinformationssysteme, Landinformationssysteme und ihre Grundlagen Springer Verlag Heidelberg, ISBN 0-387-50410-9	Lehrbuch mit den Schwerpunkten geometrische, thematische Modellierung, Methoden zur Verwaltung raumbezogener Daten, geometrische Algorithmen. *Für GIS - Kernthemen noch gut verwendbar, bei EDV-Themen nicht mehr aktuell; mit Glossar.*
Bartelme, N. (Graz 2005) **Geoinformatik** – Modelle, Strukturen und Funktionen, 3. Auflage Springer Verlag Heidelberg, ISBN 3-540-20254-4	Lehrbuch mit dem Schwerpunkt Modellierung und Verarbeitung von Geoinformation in Datenbanken in Form von Geometrie- und Rasterdaten. *Hinweise auf Normungs- und Standardisierungsaktivitäten für interoperable Geodatenmodelle.*
Bill, R. (Rostock 2010) **Grundlagen der Geo-Informationssysteme** 5. Auflage Herbert Wichmann, Hüthig Fachverlage, Heidelberg, ISBN 3-87907-489-1	Standardwerk der Geoinformatik mit Basiswissen über die Erfassung raumbezogener Daten, ihre geometrische und topologische Modellierung sowie mathematische und algorithmische Grundlagen für gängige Analysemethoden und die Präsentation von Geodaten. Überblick über das Anwendungsspektrum von GIS und neue Entwicklungen. *Das Lehrbuch ist wohl das am häufigsten zitierte deutsche GIS – Standardwerk (früher in 2 Bänden).*
De Lange, N. (Osnabrück 2005) **Geoinformatik in Theorie und Praxis** Springer Verlag Heidelberg, ISBN 3-540-28291-2	Lehrbuch mit einer Gesamtdarstellung von Geoinformationssystemen – Erfassung, Verwaltung, Modellierung, Analyse und Präsentation von Geodaten. *Das Lehrbuch enthält eine kompakte Darstellung aller funktionalen Komponenten (EVAP) von GIS.*
Göpfert, W. (Darmstadt 2002) **Raumbezogene Informationssysteme** – Datenerfassung, Verarbeitung, Integration Herbert Wichmann, Hüthig Fachverlage, Heidelberg, ISBN 3-87907-165-9	Lehrbuchartige Darstellung mit dem Schwerpunkt Rasterdaten und Bildanalyse in Kombination mit Vektordaten für den Einsatz in Raumordnung, Landes-planung und kommunaler Planung unter Ausschöpfung des thematischen Informationspotentials digitale Bilder. *Schwerpunkt ist die Verarbeitung von hybriden Geodaten.*
Hake/Grünreich/Meng (Hannover 2002) **Kartographie** – Visualisierung raum-zeitlicher Information Verlag de Gruyter, Berlin, ISBN 3-11-016404-3	Standardwerk der Kartographie mit den allgemeinen Grundlagen, der kartographischen Modellbildung, der grafischen Datenverarbeitung und kartographischen Visualisierung von Geodaten. *Viele Querbezüge und Gemeinsamkeiten zwischen Geoinformatik und digitaler Kartographie.*

Zeitschriften

GeoBIT/GIS, Herbert Wichmann, Hüthig Fachverlage (monatlich) http://www.geopoint.de	Anwenderberichte, Fachbeiträge sowie Nachrichten zu Produkten und Marktgeschehen sowie Sonderausgaben für den interdisziplinären Austausch zwischen den Geowissenschaften.
geoworld, Adams Business Media (monatlich) http://www.geoplace.com	Internationales Magazin für GIS; Anwendungen, Projekte und Softwareentwicklungen.

Internet-Adressen

http://www.geoinformatik.uni-rostock.de	***Bill:*** Geoinformatik-Service mit GI-Lexikon, Softwareprodukte, Literatur und Links
http://www.gis.bv.tum.de	***Schilcher:*** Vorlesungen (Skriptum und Folien), umfangreiches GIS-Glossar
http://ifgi.uni-muenster.de	***Streit:*** Vorlesungen, GI-Forum
http://www.igf.uni-osnabrueck.de	***de Lange:*** Vorlesungen, OGC
http://www.geolist.eu	Protal: Vermessung und Geoinformation im Internet
http://www.geobranchen.de	Protal: Links, Stellenmarkt, Daten und Software
http://www.giswiki.org/wiki/GIS-Lexikon	freies Portal für Geoinformatik (GIS)
http://www.geodaten.bayern.de	Geobasisdaten der Bayer. Vermessungsverwaltung
http://www.gdi.bayern.de	Geodateninfrastruktur Bayern (GDI-BY)
http:// www.opengeospatial.org	Open Geospatial Consortium (OGC)
http://inspire.jrc.ec.europa.eu	Geodateninfrastruktur Europa (INSPIRE)

Abkürzungen

AAA	AFIS-ALKIS-ATKIS
ACID	Atomar-Consistent-Isoliert-Dauerhaft, Konzept für Datenbanktransaktionen
AdV	Arbeitsgemeinschaft der Vermessungsverwaltungen
AFIS	Automatisiertes Festpunkt Informationssystem
ALB	Automatisiertes Liegenschaftsbuch
ALK	Automatisierte Liegenschaftskarte
ALKIS	Amtliches Liegenschaftskataster-Informations-System
ANSI	American National Standards Institute
ASCII	Amercian Standard Code for Information Interchange
ATKIS	Amtliches Topograpisch-Kartogra-phisches Informations-System
BOF	Bedienoberfläche
CAD	Computer Aided Design
CEN	Comitè Europèen de Normalisation
DB	Datenbank
DBID	Datenbank Identifier
DBMS	Datenbankmanagementsystem
DDL	Datendefinitionsprache (-language)
DFK	Digitale Flurkarte (ALK in Bayern)
DGM	Digitales Geländemodell
DHDN	Deutsches Hauptdreiecksnetz
DHHN	Deutsches Haupthöhennetz
DHM	Digitales Höhenmodell, Oberbegriff zu DGM, DOM
DIN	Deutsche Industrie-Norm
DML	Datenmanipulations- und Anfragesprache (-language)
DOM	Digitales Oberflächenmodell
DOOD	Deductive and Object-Oriented Database
DOP	Digitales Orthophoto
dpi	Dots per Inch
DTK	Digitale topographische Karte
DXF	Data Exchange Format (AutoCAD)
EDV	elektronische Datenverarbeitung
ERM	Entity-Relationship-Modell
ETRS89	European Terrestrical Reference System 1989
FFG	Fehlerfortpflanzungsgesetz
FID	Feature Identifier
FIG	Fédération Internationale des Géomètres
Flst.Nr.	Flurstücksnummer
GA	Grafikarbeitsplatz
GDB	Geographische Datenbank (-basis)
GDM	Graphendatenbankmodell (Netzwerk)
GDI	Geodateninfrastruktur
GDS	Geodatenserver
GeoTIFF	Georeferenziertes TIFF
GeoIT	Geoinformationstechnologie
GI	Geoinformatik

GIS	Geoinformationssystem, Synonyme: Geographisches Informationssystem, Raumbezogenes Informationssystem
GK	Gauß-Krüger-System
GML	Geographic Markup Language (OGC)
GNSS	Global Navigation Satellite Systems
GPS	Global Positioning System
GRS80	Geodetic Reference System 1980
GRUBIS	Amtliches Grundstücks- und Bodeninformationssystem
ISO	Internatinal Organization for Standardization
IT	Informationstechnik
JPEG	Joint Photographic Experts Group
KML	Keyhole Markup Language (Google)
LAN	Local Area Network
LIS	Landinformationssystem
MERKIS	Maßstabsorientierte Raumbezugsebenen für kommunale GIS
OFA	Oracle Flexible Architecture
OGC	Open Geospatial Consortium
OID	Object Identifier, Objektschlüssel
OMT	Object Modeling Technique
OOA	objektorientierte Systemanalyse
OOD	objektorientierter Systementwurf
OODB	objektorientierte Datenbank
OODBMS	objektorientiertes DBMS
OODM	Objektorientiertes Datenbankmodell
ORDBMS	objektrelationales DBMS
PDF	Portable Document Format (Adobe)
PNG	Portable Network Graphics (W3C)
QM	Qualitätsmanagement
QMS	Qualitätsmanagementsystem
RAID	Redundant Array of Inexpensive Disks
REO	Raumbezogenes Elementarobjekt
RDM	Relationales Datenbankmodell
ROK	Raumordnungskataster
SGA	System Global Area (Oracle)
SHAPE	Datenformat von ArcView
SICAD	Siemens Computer Aided Design
SID	System Identifier (Oracle)
SQL	Structured Query Language
SVG	Scalable Vector Graphics (W3C)
TIFF	Tagged Image File Format
TIS	Topographisches Informationssystem
TIN	Triangulated Irregular Network
TK25,TK50	Topographische Karten
UML	Unified Modeling Language
URL	Uniforn Resource Locator
UTM	Universal Transversal Mercator System
WAN	Wide Area Network
WGS84	World Geodetic System 1984
WWW	World Wide Web
W3C	Word Wide Web Consortium
XML	Extensible Markup Language
ZUSO	Zusammengesetztes Objekt (ATKIS)

0 Einführung Geoinformatik

Vermessung und Geodaten

Aufgabe der Vermessung ist
die Ausmessung und Abbildung der Erdoberfläche

„Die Geodäsie will die Erdoberfläche ***messen, berechnen und zeichnen****.*
Im Wirtschaftsleben tritt sie vielfach in Erscheinung und bildet die Grundlage für Folgearbeiten. Zu nennen sind: Sicherung der Eigentumsgrenzen, Raumordnung und Planung, Bau von Siedlungen, Straßen, Eisenbahnen, Kanälen, Be- und Entwässerungen, ...“

Geodaten beschreiben Erscheinungen und Objekte der realen Welt mit ihrem Raumbezug. Der Raumbezug legt den Ort fest, wo sich eine Erscheinungsform oder ein Objekt auf der Erdoberfläche befindet.

Die Aufgabe der Vermessung ist
primär die Erfassung, Aktualisierung, Verwaltung und Verarbeitung von Geodaten sowie die Georeferenzierung von Fachdaten.

Die Geoinformatik umfasst die IT-Methoden, die für diese Aufgaben benötigt werden

Grundlegende Prozesse des Vermessungswesens

Geodatenerfassung Amtliche Vermessung (Geobasisdaten) und projektspezifische Bestandsaufnahmen	**Absteckung** Ingenieurvermessung Übertragung geplanter Objekte in die Örtlichkeit	Operative Arbeiten sind heute stark automatisiert Schwerpunkt sind die konzeptionellen Arbeiten: • Projektplanung • Messverfahren und Messanordnung festlegen • Überwachung der Arbeiten • Bewertung der Ergebnisse
↓ **Messen** **Rechnen** **Zeichnen** (Kartieren)	**Abstecken** **Rechnen** **[Planen]** ↑	
Produkte **Koordinatenverzeichnisse, Flurkarten, Thematische Darstellungen**		
Basisdaten **GIS / Geodatenbank** Fachdaten		GIS-Funktionen: • **E**rfassen, Aktualisieren • **V**erwalten • Integrieren, Kombinieren • **A**nalysieren • **P**räsentieren, Visualisieren

Kartographie und Geodaten

Die Aufgabe der Kartographie ist der Entwurf und die Herstellung kartographischer Darstellungen unter Berücksichtigung ihres Zwecks.

(Begriffsbestimmung für die Karte, Salitschew 1967)

"Karten sind verkleinerte, generalisierte, erläuterte und ***mathematisch bestimmte Abbildungen*** *der Erdoberfläche in der Ebene, welche die Verteilung, den Zustand und die Zusammenhänge der verschiedenen natürlichen und gesellschaftlichen Erscheinungen zeigen, die entsprechend dem Zweck der Karte ausgewählt und charakterisiert werden."*

Es werden unterschieden:

- Topographische Karten, sie zeigen Form, Gliederung und Bedeutung des Geländes
- Thematischen Karten, sie enthalten vorwiegend Erscheinungen nicht-topographischer Art. Bei diesen Erscheinungen handelt es sich jedoch um Dinge, die mit der Erdoberfläche in Verbindung stehen, also eine georäumliche Lage, Verteilung oder Bewegung besitzen

→ Karten sind Visualisierungen von Geodaten und georeferenzierten Fachdaten!

Die Geoinformatik umfasst die IT-Methoden, die für diese Aufgaben benötigt werden.

Die Karte ist die wichtigste Präsentationsform für Geodaten.

Karten sind themenbezogene, analoge oder digitale Abbildungen raumbezogener Sachverhalte mit den Mitteln der kartographischen Visualisierung und unter Verwendung eines festgelegten mathematischen Modells (Projektion).

(Briggs J., Peat D., 1990, Die Entdeckung des Chaos)

Die allgemeine Bedeutung von Karten kann folgendermaßen beschrieben werden: *„Karten sind anschauliche Bilder, die es dem Denken erlauben, sich auf Aspekte der Realität zu konzentrieren, die sonst allzu leicht in den Details verloren gehen. Mit einer guten Karte lernen wir Züge der Wirklichkeit zu schätzen, die uns sonst vielleicht entgangen wären, und wir können in dieser Wirklichkeit Forschungen anstellen, die uns ohne die Karte sicher nicht gelängen".*

Anwendungsbereiche „Geo"

- **Geo**däsie (Erdoberfläche **vermessen und berechnen**, Raumbezug und Form)
- **Geo**graphie (Erdoberfläche und Wechselwirkungen Erde-Mensch **beschreiben**)
- **Geo**logie (Aufbau und Entwicklung der Erde untersuchen)
- Kartographie (Erdoberfläche in Karten **abbilden**, Visualisierung)
- Stadt- und Regionalplanung
- Landwirtschaft, Forst, Ökologie, Meteorologie (Vegetation, Biosphäre)
- Verkehr, Versorgung, Entsorgung (Technische Infrastrukturen, Utilities)
- (Geo-)Marketing, eBusiness, Statistik, Logistik (Wirtschaft), Industrieanlagen

→ Viele Berufsgruppen befassen sich mit GIS ! (konkurrieren um GIS)

Wo liegen die Stärken des Vermessers ?

→ in den exakten geometrischen Operationen !

- **Messen** von Teilen der Erdoberfläche und den mit ihr verbundenen Objekten **nach Form und Lage** (Abstrahieren und Messen)
- Herstellen des Bezugs zu übergeordneten Koordinatensystems (CRS)
- Zusammenführen und homogenisieren von Geodaten aus verschiedenen Quellen
- Geometrische Auswertungen und Abbildungen
- Beurteilung der metrischen Genauigkeit

Wo liegen die Stärken des Kartographen ?

→ in der Visualisierung der Geodaten !

Sinnvolles kartographisches Visualisieren kann durch ein System nicht ersetzt sondern nur unterstützt werden!
Es ist nach wie vor das know-how von Kartographen erforderlich, um ein Thema richtig und aussagekräftig mit den Mitteln der kartographischen Visualisierung zu gestalten, so dass eine Karte entsteht, die ihren Zweck (Meng: Affordanz) erfüllt.

Digitale Kartographie hat manuelle Arbeiten stark reduziert, eröffnet aber neue Möglichkeiten Karten zu gestalten:

- Karten sind mehr Arbeitsmittel, weniger wertvolle Unikate
- Polythemen entflechten, Zusatzkarten für Sonderanwendungen erstellen
- Spezielle Themen-Varianten für verschiedene Nutzergruppen ableiten
- Präsentation von Geoinformation mittels Medien (Video, Audio, Bilder, ..)
- Viele verschiedene digitale und analoge Medien für die Präsentation (Geomedientechnik) einsetzen

Wo braucht man andere Fachleute?

→ Zur Analyse, Beschreibung und Interpretation von **Fach**themen, diese erfordern die entsprechenden **Fach**kenntnisse, z.B. von Förstern, Ökologen, E-Technikern, ...

Anmerkung: In der Vorlesung Geoinformatik bilden die Datenstrukturen für Geodaten einen Schwerpunkt, da Datenstrukturen und Algorithmen Kernthemen der Informatik sind. Für die Arbeit von Geoinformatikern sind die Fähigkeiten ein neues Geodatenmodell zu entwerfen oder ein bestehendes zu analysieren wichtige Grundkenntnisse, da von der Struktur der Geodaten die Auswertemöglichkeiten und die benötigten Auswertemethoden abhängen. Algorithmen werden nur methodisch behandelt, da sie i.d.R. von den verwendeten Produkten vorgegeben und nicht selbst programmiert werden. Grundkenntnisse zu geometrischen Algorithmen werden in anderen Fachvorlesungen vermittelt.

1 Grundbegriffe der Informationsverarbeitung

Geoinformatik

Die Informatik befasst sich mit Datenstrukturen und Algorithmen, die für die performante Verarbeitung und effiziente Organisation von Information - als Daten in Rechnern - (Informationssysteme) erforderlich sind.

Die Geoinformatik befasst sich dementsprechend mit der Abbildung (Abstraktion) und performanten Verarbeitung von Geoinformation in Form von Geodaten in Rechnern. Dazu sind spezielle **geometrische Datenstrukturen** für die Modellierung räumlicher Erscheinungen und geometrische Algorithmen für deren Verarbeitung nötig.

Die Abstraktions- und die Erfassungsverfahren für Geoinformation sowie die Methoden für ihre kartographische Visualisierung werden in Geoinformationssysteme integriert, sie entsprechen den üblichen vermessungstechnischen oder kartographischen Prozessen.

80% aller geschäftsrelevanten Daten haben einen Raumbezug!

Dabei sind zu unterscheiden:

Daten die Erscheinungen und Objekte des Geo-Raumes (Geoobjekte)
z. B.: Flurstücke, Gebäude, Biotope, Naturschutzgebiete, Leitungen, ...
geometrisch und attributiv beschreiben, ihre **Lage, Form** und **Bedeutung** festlegen.

→ Geodaten

Daten über nicht geometrische Erscheinungen,
z. B.: Einwohner, Kunden, Geldautomaten, Hotels, Geschäfte, ...
die mit einem Ort (Koordinate) direkt oder indirekt verknüpft werden können,
zur Lokalisierung oder zum Erfassen ihrer räumlichen Verteilung (Karte mit „Fähnchen“)

→ georeferenzierte Fachdaten

	Entwicklung der Geoinformationssysteme
	Erste Entwicklungen (Pioniere) Technische Voraussetzung: Vektorgrafik
1963 1965	Canada Geograohic Information (CGIS) Synagraphics Mapping System (SYMAP)
	Erste Firmen
1969	ESRI, Jack Dangermond (Enviromental System Research Institut)
	INTERGRAPH, Jim Meadlock
1970	ALK-Sollkonzept verabschiedet
1972	Start des ersten Landsat-Satelliten
	Erste Softwareprodukte
1979 1981	Markteinführung von SICAD, Siemens Markteinführung von ArcInfo, ESRI
1980	GIS-Projekte in der Kommunalverwaltung (Großstädte) und bei Energieversorgern
1982	FIG definiert Landinformationssysteme
1984	Erste GIS-Seminare an der Universtiät Edinburgh
1985	GPS wird operational
1986	GIS-Standardwerk P. Burrough: „Principles of Geographic Information“
1988	Erste GIS-Seminare an deutschsprachigen Hochschulen
1989	Beginn der ATKIS-Einführung
1991	GIS-Standardwerk Bill/Fritsch: „Grundlagen der Geoinformationssysteme“
1993	Martkeinführung von ArcView GIS, ESRI
1994	ISO/TC 211, GIS-Normung
	OGC-Gründung, GIS-Standardisierung Ziel: Interoperable GIS
1995	GIS als eigenständiges Lehrgebiet an deutschen Hochschulen
1996	ORACLE führt die Spatial Data Option ein, DBMS-Erweiterung für Geodaten
2000	Markteinführung von ArcGIS, ESRI
2001	Einführung der Geodateninfrastruktur, Beschluss des Bundestags (GDI)

Information

Das Wort Information gilt in der deutschen Umgangssprache als Synonym von Wissensvermittlung, insbesondere im Sinne **neuen und praktisch relevanten** Wissens. Für diesen zentrale Begriff der Informationstechnik gibt es keine einheitliche Definition.

Kybernetik: *Information ist eine räumliche oder zeitliche Folge physikalischer Signale, die mit bestimmten Wahrscheinlichkeiten oder Häufigkeiten auftreten und bei einem Empfänger ein bestimmtes Denkverhalten bewirken. Informationen aus voneinander unabhängigen Quellen wirken additiv.*

Nachrichtentechnik: *Information ist ein allgemeiner Begriff für eine Nachricht, die von einer Nachrichtenquelle ausgesandt und über einen Nachrichtenkanal an Empfänger übertragen wird. Es handelt sich dabei um diskrete (digitale) oder kontinuierliche (analoge) Zeichenfolgen.*

Journalistik: *Medien vermitteln Informationen, die im Hinblick auf den Empfänger unterschieden werden, es können Neuigkeiten, Kommentare, Hintergründe, etc. sein.*

Informatik: *Information ist das, was ein Mensch als nutzbares Ergebnis aus einen Informationssystem herausholt. Sie ist eine logisch in sich geschlossene Einheit und stellt sich als höhere Ordnung den Daten gegenüber, aus denen sie sich zusammensetzt.*

Aus den unterschiedlichen Ansätzen zeigt sich, dass Information durch die Interpretation von Daten und Nachrichten gewonnen wird. Dabei kann ein und derselbe Datenbestand unterschiedlich interpretiert, verschiedene Information liefern.

Der Grundbegriff Information soll hier folgendermaßen weiter verwendet werden:

Information ist das nutzbare Ergebnis, das ein Mensch als Antwort auf seine Fragen aus einem Informationssystem herausbekommt.

Information ist damit das Resultat von Selektionen, Analysen und Synthesen der Dateninhalte eines Informationssystems, entsprechend den Fragestellungen der Nutzer.

Die Repräsentation von Information ist nicht an Daten gebunden (Beispiele)

Natur	Elektronenschale	Atome	chemische Reaktionen
	DNS-Moleküle, Erbsubstanz	Zellen	Bio-Baupläne
Sinne	Bild	Visuelle 2D-Muster	Fotos, Karten
	Schrift	Zeichenfolgen	Texte (Schriftsprache)
	Töne	Schallwellen	Sprache (Lautsprache)
Technik	Elektromagnetische Signale	modulierte elektromagnetische Wellen	Telekommunikation
	Daten	digitale Zeichenfolgen	Informations-Verarbeitung und -Speicherung im Computer

Daten

Damit Information ausgetauscht, verarbeitet und gespeichert werden kann, bedarf es einer Repräsentation, mit der die Information in interpretierbarer Form abgebildet werden kann. Die Natur verwendet dafür atomare und molekulare Strukturen (z.B. Gene, die wohl leistungsfähigsten Informationsträger). Papier und Schrift sind in der Geschichte der Menschheit der bisher bedeutendste künstliche Informationsträger. Zunehmend wichtiger werden Informationsträger, die von EDV-Anlagen verarbeitet werden können, dies sind magnetische oder optische Speichermedien mit Daten.

Die klassische *"von-Neumann-Rechnerarchitektur"*, an denen sich die heutigen Anlagen noch immer orientieren, umfasst insbesondere den Prozessor, dessen Operationen Speicherzellen bearbeiten. Jeder Informationsverarbeitungsprozess muss also letztlich in Speicherinhalten oder Daten und auf diese wirkende Folgen von Maschinenoperationen (Algorithmen) formuliert werden. Unter Daten versteht man in der EDV daher alles, was sich in einer für Rechner erkennbaren Weise kodieren lässt.

Die unterste Kodierungsebene für Daten in einem Rechner ist das Bit (binary digit), die kleinste darstellbare Informationseinheit, mit der die Prozessoren arbeiten. Alle Abläufe in einem Prozessor müssen auf binäre Operationen zurückgeführt werden.

Aus Anwendersicht sind jedoch höherwertige Datenaggregate erforderlich, um die Abbildung von Information in einem Rechner zu ermöglichen. Dies sind im einfachsten Fall Zahlen (Real/Integer) oder Zeichenketten (Charakter) aus denen wiederum komplexere Strukturen gebildet werden können.

Datenaggregate

Kleinste Kodierungseinheit	BIT	Prozessor
Zeichen	BYTE, ASCII-Zeichensatz	Benutzer
Datentypen	Character (String, Wort) Integer, Real (Zahl) Logical, Pixel	elementare Datentypen für die Programmierung
Abstrakte Datentypen (Datenstrukturen)	Attribut-Tupel (Satz) Geometrieelement Bildmatrix (Feld)	komplexe Datentypen und zugehörige Operationen für die Modellierung
Datenmodelle	OODBM: Objekte (Klassen) RDBM: Relationen	Spezifizierte, kombinierte und strukturierte Daten

Daten sind die Repräsentation der Information in IT-Systemen,
sie können von Prozessoren verarbeitet werden und
sind auf digitale und analoge Speichermedien eindeutig abbildbar.

Daten sind die Grundlage der Informationsverarbeitung in IT-Systemen, sie sind passiv-statisch und werden erst durch ihre Interpretation, im Kontext einer konkreten Fragestellung, wieder zur Information. Das Datenvolumen sagt a priori nichts über ihren Informationsgehalt aus. Ein großer Datenbestand kann nur aus zufälligen Zeichen (Rauschen) bestehen und somit keine Information beinhalten.

Unterschied zwischen Information und Daten

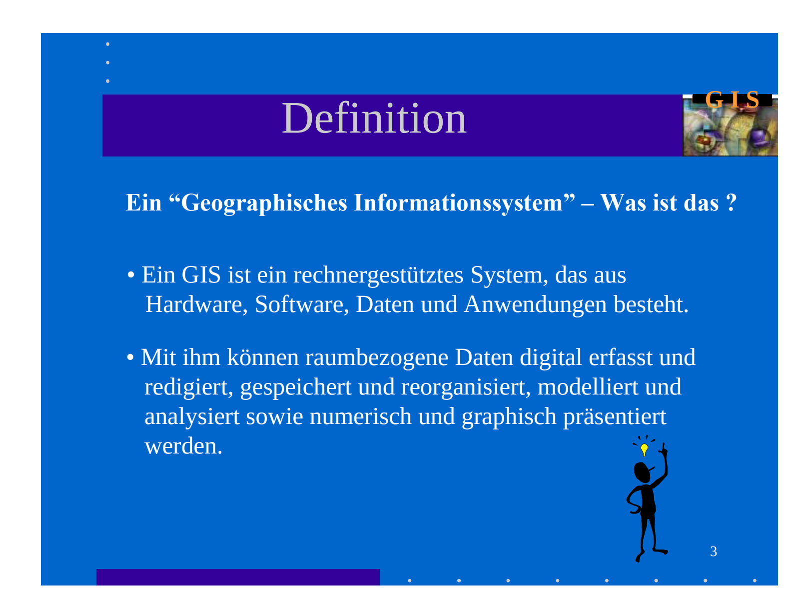
Rasterbild

Für die Repräsentation eines Bildes mittels Rasterdaten in einer EDV-Anlage werden gleichförmige rechteckige oder quadratische Bildelemente (Pixel) verwendet. Die Bildelemente sind nach Zeilen und Spalten in einer Matrix angeordnet. Zu jedem Bildelement wird ein Farb-, Grau- oder Binärwert gespeichert, zwischen den Bildelementen bestehen keine logischen Beziehungen. Untersucht man die verschiedenfarbigen Bildpunkte einzeln, lässt sich der Bildinhalt - seine Information - nicht ohne weiteres erkennen. Die analytische Betrachtung der Rasterdaten, z.B. in Form eines Histogramms der Grauwerte oder deren Mittelwerte führt i.d.R. zu keiner Lösung.

Erst die synthetische Gesamtsicht aller Einzeldaten, insbesondere ihre räumliche Anordnung, lässt die Information erkennen.

(Frederic Vester, 1985, Neuland des Denkens)





PowerPoint-Folie	215 kB
Hintergrund	200 kB
Textfelder und Steuerzeichen	15 kB
Relevante Zeichen	~ 0,5 kB
Informationsgehalt	????

$I(A) = -\log P(A)$ *mit* $P \in [0,1]$

Ist abhängig vom Wissen des Betrachters,
also subjektiv !

Joseph Weizenbaum *(2000, sagt in einem Interview zum Thema Internet)*

„Weder der Computer noch das Internet enthalten Wissen, sondern nur Daten. Erst wenn ein Mensch diese Daten interpretiert und in einen Zusammenhang einbaut werden sie zur Information. Wenn diese Information vom Menschen vernünftig verarbeitet wird, führt sie langsam zu Wissen.“

Informationssystem Datenbank	Kommunikation		Benutzer Hintergrundwissen
	⇐ Frage, Auswahl ⇐		
Daten ⇒ <Informations-potential>	Ergebnis Daten	⇐ Interpretation Information ⇒	⇐ **Wissen** Neues Wissen

Nachrichten

Für die Kommunikation zwischen Mensch und Computer ist es nötig, dass die Fragen und Anweisungen des Anwenders der Maschine mitgeteilt werden können und die Maschine zu einem hinreichenden Dialog mit dem Benutzer fähig ist. Diese Wechselbeziehung bezeichnet man heute als Interaktion, dabei wechseln sich Aktionen des Benutzers mit Reaktionen des Systems ab.

Die Interaktion mit einem Informationssystem besteht im Austausch von Nachrichten, die übermittelt werden, um die Informationsverarbeitungsprozesse zu steuern. Auch die Komponenten eines Informationssystems kommunizieren untereinander mit Nachrichten.

Daten und Nachrichten sind beide Folgen von Zeichen, sie unterscheiden sich jedoch in ihren Zielsetzungen.

Während Daten der Verarbeitung und Speicherung von Information in IT-Systemen dienen, ist der Zweck von Nachrichten die Informationsübertragung.

Daten und Nachrichten bedürfen der Interpretation, um ihren Inhalt und Zweck zu erfahren. Die Mechanismen (Algorithmen) dazu enthalten die Vorschriften nach denen die Zeichen interpretiert werden, dabei sind drei Ebenen zu unterscheiden:

Syntax: die Ebene der **Zeichen**, sie regelt die formale Relation der Zeichen (Reihung, Form, Größe, Abstand, Kontrast, ...)

Semantik: die Ebene der **Bedeutungen**, sie regelt die Beziehungen zwischen den Zeichen und den Objekten, die sie beschreiben

Pragmatik: die Ebene der **Beziehungen**, sie regelt den Zusammenhang zwischen den Zeichen und den Benutzern
(z.B. Zweck und Empfänger einer Nachricht)

Entitäten und Objekte

Eine Entität ist eine **eindeutig identifizierbare Erscheinung oder ein Ereignis in der realen Welt**, sie kann unabhängig von anderen Entitäten existieren, und sie lässt sich in Phänomene anderer Art, jedoch nicht in solche der gleichen Art unterteilen.

Unter einem **Objekt** wird hier ein allgemeines technisches Konzept verstanden, mit dem man die realen Objekte der Umwelt und deren Beziehungen, die hier zur Unterscheidung als Entitäten bezeichnet werden, so abstrahieren kann, dass sie von einem Rechner als logische Informationseinheiten verarbeitet und verwaltet werden können.

Entität (reale Welt) → Abstraktion → **Objekt (virtuelle Welt)**

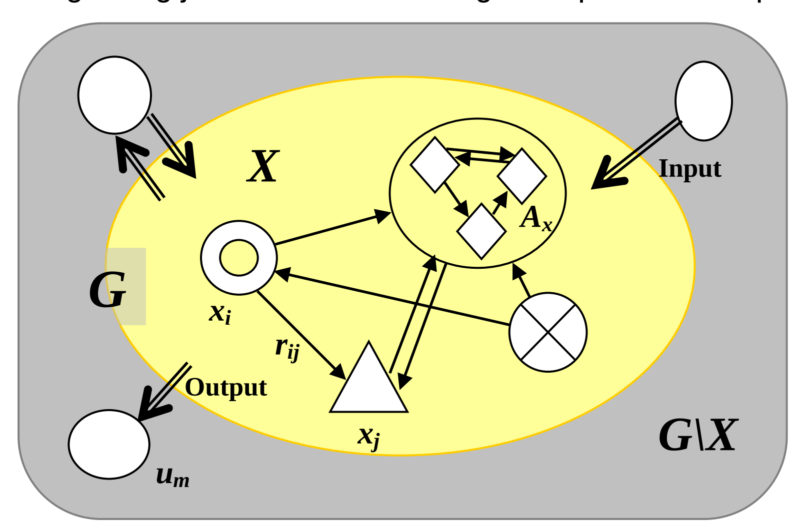
„Entitäten sind die Gegenstände der Informationsverarbeitung“

Die nötige Abstraktion bei der Abbildung von Entitäten in einem Informationssystem aus Sicht der jeweiligen Fachanwendung führt zu einer mehr oder weniger starken Generalisierung der Entitäten, d.h. die Reduktion auf ihre charakteristischen Eigenschaften im Fokus des jeweiligen Fachgebietes.

Systeme

In der *DIN-Norm 19226* wird ein System beschrieben als *”... eine Menge aus einer Anzahl von definierten Elementen, wobei für jedes Element die Entscheidung über seine Zugehörigkeit zur Menge eindeutig sein muss. Es bestehen zwischen Elementen unterschiedliche Beziehungen dergestalt, dass jedes Element durch eine oder mehrere Beziehungen direkt oder indirekt verbunden ist.”* Der Kern des Systembegriffes ist, dass ein System Eigenschaften hat, die nicht aus einer bloßen Reihung der Eigenschaften seiner Elemente erklärt werden können, sondern die durch das Beziehungsmuster der Elemente determiniert sind.

Definition des Systembegriffs: Ein System $S=\{X,R\}$ ist gegeben, wenn eine umfassende Grundgesamtheit G eine kleinere Trägermenge X umschließt, die sich aus kleinsten Einheiten x_i aufbaut. Diese Elemente x_i unterhalten Wirkungsbeziehungen $R=\{r_{ij}|\text{Relation der } x_i \in X\}$ untereinander und zu Einheiten, die der Umgebung $G\backslash X$ angehören; diese Relationen sind systembildend**.** Falls Elemente x_i Aggregationen in X ergeben, werden diese Aggregate A_x als Subsysteme bezeichnet. Folgende Minimalanforderungen an ein System werden gestellt: Die Trägermenge X muss mindestens zwei Elemente x_i enthalten; konkrete Einheiten der Umgebung $G\backslash X$ müssen sich von den Elementen x_i unterscheiden, aber mit ihnen interagieren; zwischen den Elementen des Systems und den Einheiten der Umgebung muss mindestens eine Relation (Wirkungsbeziehung) existieren. Bei den Aktivitäten eines Systems lassen sich endogene Aktivitäten die von den systembildenden Wirkungsbeziehungen ausgehen, und exogene Aktivitäten die von der Umgebung $G\backslash X$ verursacht werden, unterscheiden. Des Weiteren lassen sich die Wirkungsbeziehungen zwischen dem System und seiner Umgebung je nach ihrer Richtung als Input oder Output bezeichnen.



System $S=\{X,R\}$
Grundgesamtheit G
Trägermenge $X \subset G$
Umgebung $G\backslash X$
Elemente (Einheiten) $x_i \in X$
Aggregate A_x *(Subsysteme)*
Struktur $R=\{r_{ij}|\text{Relation der } x_i \in X\}$
endogene Wirkungsbeziehungen r_{ij}
Einheiten der Umgebung $u_m \in G\backslash X$
Input/Output zur Interaktion
exogene Wirkungsbeziehungen

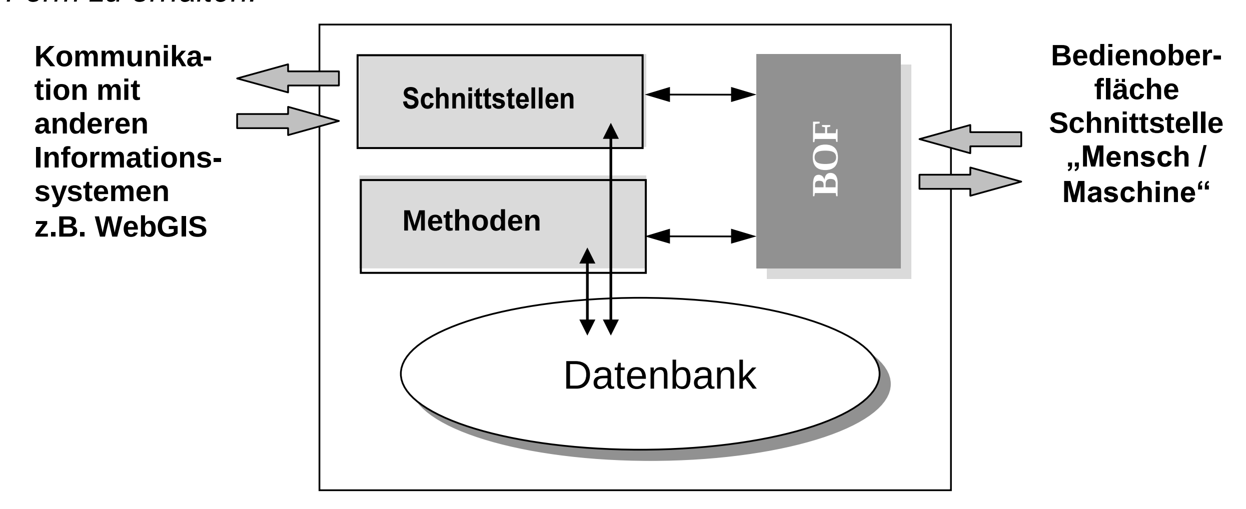
Für Systeme gilt das von *Liebig (1803-1873)* formulierte Gesetz vom Minimum, das sinngemäß lautet:

”Das schwächste Element eines Systems,
das für seinen Betrieb wesentlich ist,
bestimmt seine Leistung.”

Dieser Satz weist besonders darauf hin, dass zur Leistungsoptimierung, die Beseitigung der Schwachstellen eines Systems nötig ist, nicht die weitere Verstärkung anderer, bereits leistungsfähiger Elemente.

Informationssysteme

In seiner einfachsten Form ist ein Informationssystem ein Frage-Antwort-System auf einen Datenbestand. In *Bill/Fritsch* wird ein Informationssystem folgendermaßen definiert: *"Beschränkt sich die Funktion des Systems auf die Aufnahme (Erfassung), Verarbeitung und Wiedergabe von Information, so ist es ein Informationssystem. Es besteht somit aus der Gesamtheit der Daten und Verarbeitungsanweisungen. Der Benutzer soll in der Lage sein, daraus ableitbare Information in einer verständlichen Form zu erhalten."*



Ein Informationssystem besteht aus folgenden wesentlichen Elementen:

Datenbank – Methoden – Schnittstellen – Benutzeroberfläche

In diesem Sinne ist unter einem Informationssystem ein System zu verstehen, dessen aktive Elemente (Komponenten) Informationsverarbeitung wahrnehmen:

Erfassung (Input) – Verwaltung – Verarbeitung (Analyse) – Präsentation (Output)

und untereinander Daten austauschen und durch formalisierte Kommunikationsbeziehungen miteinander verbunden sind.

Datenbank

Eine Datenbank ist eine strukturierte Menge von Daten, die mit einem gemeinsamen Verwaltungssystem gepflegt wird und das Ziel hat, die Datenverwaltung von der Programmlogik der Verfahren und Methoden zu trennen. Aus EDV-Sicht handelt es sich dabei um Datenmengen, die nach einem vorgegebenen Schema dauerhaft (persistent) gespeichert, verwaltet und von verschiedenen, unabhängigen Anwendungen genutzt werden. Datenbanken werden durch physische und logische Prinzipien der Informationsverarbeitung gekennzeichnet. Die physischen Datenabbildungsverfahren beschreiben die Prinzipien der Informationsabbildung auf ein digitales Speichermedium. Die logischen Prinzipien der Abbildung enthalten das Konzept und die Regeln zur Abbildung der Informationseinheiten (Objekte und deren Beziehungen) und Strukturen zur Speicherung und Verwaltung der Datenmengen. Die Abbildung der Daten - das Datendesign - in einer Datenbank erfolgt auf verschiedenen Abstraktionsebenen, die Schemata genannt werden. Man unterscheidet nach *ANSI/SPARC (1975)* das

- **Externe Schema** (fachliche Sicht der Objekte, logische Teilsichten)
- **Konzeptuelle Schema** (logische Gesamtsicht)
- **Interne Schema** (physikalische Sicht)

Das konzeptuelle Schema beschreibt die gesamten logischen Beziehungen der Daten. Dabei geht es z.B. um die Zuordnung von Attributen zu Objekten und um Beziehungen zwischen Objekten. Im internen Schema werden die Art und der Aufbau der physikalischen Datenstrukturen beschrieben, z.B. mit wie vielen Bytes ein bestimmtes Attribut an welcher Stelle eines Datensatzes gespeichert werden soll und wie die Zugriffe auf das Attribut geregelt sind. Das externe Schema legt fest, wie Teile der Daten aus Benutzersicht (View) aufgabenorientiert darzustellen sind und welche Zugriffsberechtigungen der einzelne Nutzer hat.

Bei der Konzeption und dem Entwurf von Datenmodellen werden heute zunehmend vier statt drei Datenbankschemata genannt. Das konzeptuelle Schema wird weiter aufgeteilt in das eigentliche, abstrakte konzeptuelle Schema, das von der Realisierung unabhängig ist, und in das logische Schema zur konkreten Modellierung der Daten, das vom jeweilig eingesetzten Datenbanksystem abhängig ist.

Nach Dittrich (1997) ist eine Datenbank eine Menge zusammengehörender Daten mit folgenden Eigenschaften:

- Dauerhaft verfügbar (persistent bzw. mit explizit steuerbarer Lebensdauer)
- Konsistent, integer, sicher (Transaktionskonzept, Before-/Afterimage)
- Redundanzarm
- Potentiell groß (nicht absehbare Extension)
- Integriert (für mehrere Anwendungen mit überlappendem Informationsbedarf einsetzbar)
- Parallel zugreifbar (gleichzeitig mehrfachnutzbar, Nebenläufigkeit)
- Transparent, bei Verteilung im Rechnernetz
- Verwendbar unabhängig vom Erzeugungsprogramm
- Bequem, flexibel und effizient handhabbar (assoziativer Zugriff)

Die vier generischen Grundoperationen einer Datenbank sind:

Auswählen und Lesen (select)

Erfassen (insert), Ändern (update) und Löschen (delete)

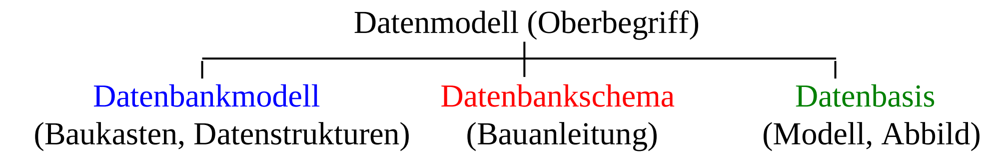
Die Verarbeitung von Operationen, die Daten verändern erfolgt in Transaktionen.

Transaktionen sind **atomare** Verarbeitungsschritte auf einer Datenbank, die immer **ganz** oder **gar nicht ausgeführt** werden müssen. Sie überführen eine Datenbank von einem konsistenten Zustand in einen anderen. Es darf keine Zwischenstadien geben, da sonst die Konsistenz der Datenbank nicht gewährleistet ist.

Transaktionen müssen die sogenannten **ACID**-Bedingungen erfüllen:

- **Atomar:** eine Transaktion ist atomar, alle Änderungen werden wirksam oder gar keine
- **Konsistent:** der gesamte Datenbestand der Anwendung ist ständig in logisch konsistentem Zustand
- **Isoliert:** Änderungen nicht abgeschlossener Transaktionen sind für andere Nutzer nicht sichtbar
- **Dauerhaft:** am Ende einer Transaktion sind die vorgenommenen Änderungen unwiderruflich

Datenmodelle



Leistung und Akzeptanz eines Informationssystems hängen wesentlich vom Informationsgehalt der Datenbank ab, d.h. von ihrem **Informationspotential**, das dem Benutzer durch die Interpretation und Auswertung der Daten erschlossen werden kann, aber auch von den Zugriffsmöglichkeiten und den Zeiten für die Informationsbereitstellung. Ein grundlegender Faktor dafür ist die optimale Adaption des Sachverhaltes durch das verwendete physische Prinzip für die Datenmodellierung.

Allgemein sind Modelle vereinfachte, zweckorientierte Beschreibungen von Teilen der realen Welt, aus fachlicher Sicht. Es sind Arbeitsmittel, die es ermöglichen, die Fülle der Umweltinformation durch sinnvolle Reduktion und Ordnung fassbar zu machen.

Voraussetzung für eine Datenmodell ist, dass es einerseits komplex genug für die Wiedergabe der in der realen Welt tatsächlich existierenden Erscheinungen und Beziehungen ist, andererseits muss es einfach genug für den effizienten und performanten Betrieb eines Informationssystems sein. Die Aufgabe der Informatik ist es nicht in erster Linie möglichst „natürliche“ Datenmodelle zu schaffen, sondern solche, die von Rechnern gut verarbeitet werden können.

Damit die Abbildung eines Interessengebiets in Form von Daten möglich ist, bedarf es der Definition eines Regelwerks (konzeptuelles Schema), das festlegt, wie die Entitäten mit Daten zu beschreiben, zu strukturieren und zu verarbeiten sind. Die Qualität der Abbildung ergibt sich aus der im Datenmodell definierten nominellen Wirklichkeit.

$$\mathbf{I}_E \xrightarrow{a} \mathbf{M} \xrightarrow{i} \mathbf{I}_A \quad \text{mit} \quad \mathbf{I}_E \geq \mathbf{I}_A$$

Die Abbildung $\mathbf{a}$ der Information $\mathbf{I}_E$ auf das Datenmodell $\mathbf{M}$ ist i.d.R. surjektiv und damit nicht eindeutig umkehrbar, da die reale Welt abstrahiert werden muss. Der Umfang und Grad der Abstraktion bei der Modellierung muss bedarfsgerecht gewählt werden. Die Interpretation $\mathbf{i}$ der Daten liefert die Ausgabeinformation $\mathbf{I}_A$. Sie kann aus verschiedenen Blickwinkeln und Zielsetzungen erfolgen, daraus ergibt sich, dass derselbe Datenbestand verschieden interpretiert, verschiedene Information liefern kann. Die Informationsmenge der Ausgabe kann die Informationsmenge der Eingabe jedoch nicht überschreiten. Durch die Kombination und Zusammenführung von Ausgabeinformation unabhängiger Datenquellen lässt sich die Informationsmenge für den Nutzer steigern, da die Teilmengen additiv sind, falls unkorrelierte Datenquellen vorliegen.

Eine Datenstruktur ist eine Organisationsform für einen Datensatz oder eine Datenmenge (z.B. Tabelle, Telefonbuch), sie soll so sein, dass bestimmte Operationen auf dieser Datenmenge effizient ausgeführt werden können.

Operationen enthalten sogenannte Algorithmen. Ein Algorithmus ist ein zusammenhängender, geordneter Satz von Regeln und Verarbeitungsvorschriften für die schrittweise Lösung eines Problems. Für die Formulierung eines Algorithmus ist die Systematik einer Aufgabe zu ermitteln und die Lösung in parametrisierbare, elementare Handlungen zu zerlegen, häufig werden iterative Verfahren eingesetzt. Wichtig ist die Optimierung der Algorithmen, um möglichst robuste und schnelle Abläufe zu erhalten. In Algorithmen können auch Zusammenhänge zwischen Daten modelliert werden, die zusätzliche Information bei der Interpretation $\mathbf{i}$ einbringen.

Für den Aufbau von Informationssystemen lässt sich strukturelles Wissen über die Umwelt in Termen von Gegenständen (Entitäten) mit ihren Eigenschaften sowie Assoziationen (Beziehungen) zwischen ihnen ausdrücken. Diese Strukturen lassen sich allgemein mit semantischen Datenmodellen (UML, ERM) beschreiben. Für den logischen Entwurf und seine Realisierung mit einem konkreten Datenbankprodukt muss jedoch das vom Produkt unterstützte Datenbankmodell eingesetzt werden.

Die wichtigsten logischen Datenbankmodelle (Datenstrukturen) sind:

- **GDBM: Netzwerkmodell (Graphen)**
- **RDBM: relationales Modell (Tabelle)**
- **OODBM: objektorientiertes Modell**

GDBM: Das Netzwerk- oder Graphenmodell ist charakterisiert durch Master-Detail-Beziehungen. Ein Master ist ein Element, dem andere Elemente (Details) nachgeordnet sind. Ein Master kann mehrere Details haben, ein Detail kann mehreren Mastern zugeordnet sein. Die Zuordnung erfolgt in Sets, jeder Set-Typ besitzt einen Master- und einen Detailtyp, womit auch komplexe Beziehungen (m:n) optimal abgebildet werden können. Das Netzwerk ist wegen seiner engen Verwandtschaft mit Graphen besonders gut für die Abbildung von geometrischen und topologischen Sachverhalten geeignet, da deren ”natürliche” Beziehungsstruktur ein Netzwerk ist. Das hierarchische Datenbankmodell ist ein Spezialfall des Netzwerks, in dem es nur (1:n)-Beziehungen gibt.

RDBM: In relationalen Modellen *(Codd 1970)* ist die einzige Datenstruktur die Tabelle. Zur Beschreibung der Objekte werden gleichberechtigte Tabellen aufgebaut, deren Zeilen (Tupel) ein konkretes Objekt mittels der festgelegten Attribute (Spalten) und ihren zugeordneten Wertebereichen (Domänen) beschreiben. Die Stärke des relationalen Modells liegt in der lexikalischen Beschreibung von Objekten mittels Attributen, die als atomare, unstrukturierte Einheiten betrachtet werden, und der ihnen zugrundeliegenden klaren mathematischen Darstellung (Cantorsche Mengen). Die Beziehungen (Relationship) zwischen den Objekten werden ebenfalls durch Attribute realisiert. Die Abbildung von komplexen Objekten, wie sie in geometrischen Strukturen die Regel sind, bereitet den relationalen Datenbanken jedoch Probleme.

OODBM: Wegen der Probleme bei der Abbildung komplexer Objekte in relationalen Datenbanken, hat sich die Forschung im letzten Jahrzehnt verstärkt mit einem weiteren Modelltyp, der objektorientierten Datenbank befasst. Eine objektorientierte Datenbank präsentiert sich dem Benutzer in Form persistenter (persistent = andauernd, anhaltend) Objekte, gleich wie sich die relationale Datenbank in Form von Tabellen zeigt. *Atkinson 1989*: *”Ein objektorientiertes Datenbanksystem dient der Verwaltung persistenter Objekte. Seine Basisfunktionalität unterstützt den Lebenslauf persistenter Objekte, d.h. das Erzeugen, das Deaktivieren, das Aktivieren und das Zerstören von persistenten Objekten.”* Darüber hinaus muss es alle Eigenschaften eines klassischen Datenbanksystems haben. Die konkrete Ausprägung eines Objektes bezeichnet man als Instanz der zugeordneten (Objekt-)Klasse. Objekte können in einer ”Teile-Beziehung” stehen und zusammengefasste (komplexe) Objekte bilden.

Derzeit gibt es zwei Trends zur Realisierung objektorientierter Datenbanken. Die eine Richtung versucht die etablierten relationalen Datenbanken zu erweitern die andere entwickelt völlig neue Datenbanksysteme

objektrelationale DB = ORDB ←→ objektorientierte DB = OODB

Methoden

Methoden (Operationen, Funktionen) und Verfahren sind die Werkzeuge, deren sich der Benutzer für die Selektion, Analyse und Darstellung der Daten bedient. Je nach Anwendungsbereich eines Informationssystems sind unterschiedliche Methoden erforderlich. Beispiele sind die Selektion, Präsentation (grafisch oder tabellarisch), Transformation, oder Verschneidung sowie statistische Auswertungen oder komplexe Simulationen.

Als Methoden werden Operationen und Funktionen bezeichnet, wenn sie als Implementierungen in Objekten bereitgestellt werden.

Man unterteilt die Methoden in:

- Konstruktoren / Destruktoren zum Erzeugen und Löschen von Objekten
- Mutatoren zum Ändern von Objekten
- Akzessoren für den lesenden und schreibenden Zugriff auf Attributwerte
- Operationen und Funktionen (Tätigkeiten)

Die Begriffe Operation und Funktion werden häufig synonym verwendet, man kann sie jedoch unterscheiden in:

- Operation, wenn Mengen von Daten verknüpft werden, das Ergebnis einer Operation ist wieder eine Menge von Daten; Beispiele sind die Verschneidung oder die Aggregation von Daten.
- Funktion, wenn aus Daten ein Merkmalswert ermittelt wird, das Ergebnis einer Funktion ist ein Wert; Beispiele sind die Flächen- und Längenberechnung.

Fasst man solche Methoden zu formalisierten Abläufen zusammen, die für die Abwicklung von standardisierten Aufgaben eingesetzt werden, so spricht man von Verfahren.

Kommunikation / Interaktion

Bedienoberfläche: Als Bedienoberfläche (BOF) eines Informationssystems versteht man zum einen die sichtbaren Eigenschaften, die die Kommunikation zwischen System und Benutzer bestimmen und zum anderen, in einem umfassenden Verständnis, die Hilfsmittel für die Interaktion des Benutzers mit dem System. Die BOF ermöglicht den Benutzer den Dialog mit dem Informationssystem, sie ist das Bindeglied zwischen Mensch und Maschine, die Dialogschnittstelle. Die Möglichkeiten und Form der Kommunikation ergeben sich aus den verfügbaren Eingabeelementen, Beispiele sind: Kommandosprachen, grafische Oberflächen (GUI) oder die Spracheingabe. Allen gemeinsam ist, dass über das Eingabemedium BOF der Maschine mitgeteilt wird, was sie tun soll. Die Funktionen der BOF beeinflussen die Akzeptanz eines Informationssystems wesentlich.

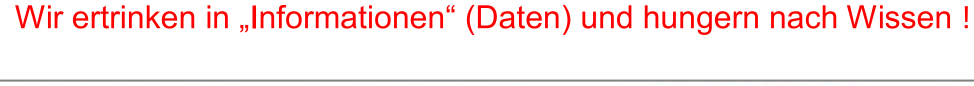
Schnittstellen (Datenschnittstellen): Unter Schnittstellen (Interface) versteht man in der EDV die Übergänge (Nahtstellen) zwischen verschiedenen Hard- und Softwarekomponenten und den Datenbeständen des EDV-Systems. Die Hardwareschnittstellen für Informationssysteme unterscheiden sich nicht von denen anderer IT-Systeme. Die verfügbaren Softwareschnittstellen bestimmen die Skalierbarkeit des Systems im Hinblick auf seine Funktionalität (gebräuchlich sind Schnittstellen für Programmierung und Makrobildung, Skripts). Schnittstellen für den Datenaustausch sind in Informationssystemen besonders wichtig, da über sie die Kommunikation mit anderen IT-Systemen und damit die Integration des Informationssystems in eine übergeordnete IT-Umgebung und die fachübergreifende Zusammenarbeit ermöglicht wird. Die Datenschnittstellen (Transfermodelle) enthalten die Vorschriften für die Übertragung von Teilen (Projektionen) des Modells von einem Quellsystem zu einem Zielsystem, sie bestimmen das Kommunikationsniveau zwischen den Systemen.

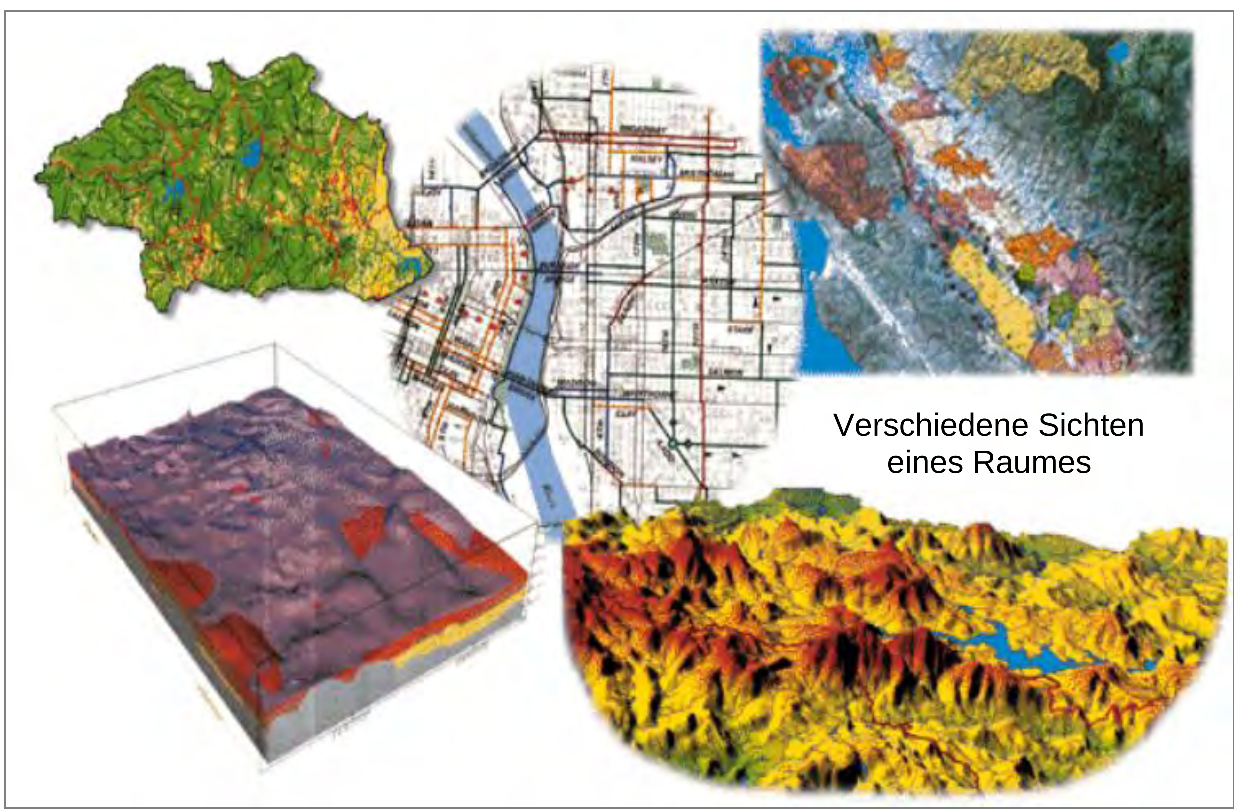
2 Geoinformationssysteme

Begriffsbestimmung und Definition

Der Bedarf an Informationen über unseren Lebensraum ist heute sprunghaft gestiegen. Hauptursachen sind die hohe Siedlungsdichte, unser Streben nach wirtschaftlichem Wachstum, steigendem Lebensstandard und höherem Freizeitangebot. Immer häufiger führen raumbedeutsame Maßnahmen zu konkurrierenden Nutzungsansprüchen an Grund und Boden und zu Konflikten mit der Ökologie.

Eine optimale Gestaltung, nachhaltige Bewirtschaftung und Gesamtentwicklung des anthropogenen Lebensraumes, im Einklang mit der Ökologie, muss daher ein primäres Ziel bei allen relevanten Planungen und Maßnahmen sein. Damit dieses Ziel erreicht werden kann, ist die Abstimmung der beteiligten Planungsträger und eine sorgfältige fachübergreifende Planung erforderlich





Informationssysteme, deren Schwerpunkt es ist, Daten über unseren Lebensraum zu verarbeiten, werden heute oft unter dem Begriff Geographische Informationssysteme zusammengefasst. Die daneben häufig benutzte Bezeichnung

Geoinformationssysteme (GIS)

benennt diese Systeme jedoch treffender und allgemeingültiger, denn nicht nur die Geographen, sondern alle Fachdisziplinen, die Geoinformation verarbeiten, bedienen sich zunehmend dieser Instrumente. Geographische Informationssysteme sind spezielle GIS, die für kleinmaßstäbliche Anwendungen in der Geographie eingesetzt werden.

Es gibt eine ganze Reihe von Definitionen und Begriffsbestimmungen für Geoinformationssysteme (GIS). *Bill/Fritsch (1991)* definieren:

"Ein Geoinformationssystem ist ein rechnergestütztes System, das aus
H*ard- und* ***S****oftware,* ***D****aten und den* ***A****nwendungen besteht.*
Mit ihm können raumbezogene Daten

- *digital erfasst und redigiert* (***E****rfassung),*
- *gespeichert und reorganisiert* (***V****erwaltung),*
- *modelliert und analysiert sowie* (***A****nalyse)*
- *alphanumerisch und graphisch präsentiert werden."* (***P****räsentation)*

Hier soll für ein GIS eine weitere Definition verwendet werden, die der Kommunalverwaltung entstammt und sich an die ursprüngliche Definition eines Landinformationssystems *(FIG 1981)* anlehnt:

"Ein GIS ist ein Instrument in
Gesellschaft, Politik, Verwaltung, Recht und Wirtschaft
für die Dokumentation, Planung und Entscheidungsfindung
bei Sachverhalten, die auf Grund und Boden - den Raum - bezogen sind.

Es besteht aus:

- ***einer geographischen Datenbank, in deren Datenbasen die Modelle raumbedeutsamer Strukturen einer bestimmten Region dokumentiert sind***
- ***Verfahren und Methoden, mit denen die Modelldaten erfasst, aktualisiert, präsentiert, assoziiert und analysiert werden können***
- ***Schnittstellen für die Integration und die Kommunikation mit anderen IT-Systemen***

Die Basis bildet ein einheitliches räumliches Bezugssystem,
das die eindeutige räumliche Zuordnung und fachübergreifende Verknüpfung der raumbezogenen Daten ermöglicht."

Diese Definition sagt sowohl etwas über die Einsatzbereiche eines GIS, als auch über seine wesentlichen Komponenten aus und beschreibt ein GIS als eine

Spezialisierung herkömmlicher Informationssysteme
für mehrdimensionale, geometrieorientierte Daten**,** die sogenannten **Geodaten**.

Raumbezug, das charakteristische Merkmal von Geodaten

Die metrische Basis eines Geodatenmodells wird als einheitliches Raumbezugssystem bezeichnet, das die eindeutige geographische Zuordnung und die fachübergreifende Verknüpfung der Geodaten ermöglicht. Die mathematische Festlegung des Raumbezugs erfolgt durch die Wahl eines geodätischen Bezugssystems, das einem GIS zu Grunde gelegt wird. Das geodätische Bezugssystem dient primär der Koordinatengebung (Geokodierung) zur Abbildung der Erdoberfläche. Dazu wird das Bezugssystem mittels physikalischer und geometrischer Festlegungen theoretisch beschrieben; diese Festlegungen werden als geodätisches Bezugssystem bezeichnet.

Für die Zahlengebung, dem Geokode, muss noch ein geeignetes Koordinatensystem zugeordnet werden. Als Koordinatensystem werden in GIS ebene Projektionskoordinaten bevorzugt, die räumliche Analyse setzt idR. kartesische Koordinaten voraus. Für die Nutzung eines Bezugssystems in einem GIS oder für geodätische Zwecke ist seine „Materialisierung“ in Form eines sogenannten Bezugsrahmens (Frame) erforderlich.

Bezugsrahmen werden üblicherweise durch ein Netz diskreter Anschlusspunkte realisiert, dies können abgemarkte Punkte im Feld oder vorgegebene virtuelle Punkte in der Datenbank sein. Diese Verbindungselemente müssen geeignet sein, die Anbindung an das Raumbezugssystem für die Geocodierung der Objekte hinreichend genau zu ermöglichen. Wegen der besonderen Bedeutung des Raumbezugs für ein GIS wird im deutschsprachigen Bereich, vor allem in der Verwaltung, auch der Begriff

Raumbezogene Informationssysteme

verwendet. Raumbezogen ist hier im Sinne des Raumordnungsgesetzes zu verstehen, d.h. es handelt sich um Sachverhalte, die mit dem Grund und Boden, dem GeoRaum zu tun haben. Der Begriff Raum weist zudem auf die wichtigste Eigenschaft der Geodaten, ihre mehrdimensionale geometrische Struktur direkt hin. Im Englischen hat sich für Geodaten der Begriff "**spatial data**" (spatial = räumlich) etabliert.

Der Raumbezug ist das Basisverknüpfungskriterium in einem GIS, für das es in einem normalen Informationssystem keinen Vergleich gibt. Ein GIS ist ein Instrument zur Umsetzung räumlichen Denkens

→ ESRI Solgan: „Thinking spatially“

Auf folgende Fragestellungen kann ein Geoinformationssystem Antworten geben:

- Was gibt es an einem bestimmten Ort? (lokale Frage)
- Welcher Ort erfüllt bestimmte Kriterien? (konditionale Frage)
- Welche Objekte stehen in Beziehung zueinander? (assoziative Frage)
- Was passiert bei bestimmten Bedingungen? (modellbasierte Frage)

Wichtige Merkmale eines GIS sind die Verwendung von Geometriedaten in Kombination mit beschreibenden Attributen und die Möglichkeit, Daten aus verschiedenen Themenbereichen, die denselben Raumbezug verwenden, zu überlagern. Für die Überlagerung sind keine zusätzlichen semantischen Verknüpfungen nötig (aber ratsam, wegen der begrenzten metrischen Genauigkeit = Unschärfe).

Anmerkung zur „Maßstabsfreiheit“ von Geodaten: Die Datenerfassung im GIS ist maßstabsorientiert, da die Geoobjekte in Abhängigkeit von ihrem Erfassungsmaßstab vereinfacht sind oder werden; ihre Form unterliegt einer Erfassungsgeneralisierung (Abstraktion). Durch die Wahl der Auflösung (Punktdichte / Genauigkeit) bei der Datenerfassung wird auch ein Maßstabsbereich festgelegt, der die erreichbare Auswertegenauigkeit implizit bestimmt. Im Hinblick auf signifikante metrische Auswertungen und Verschneidungen muss a priori auf eine hinreichende Dichte und Genauigkeit der Stützpunktinformation des Geodatenmodells geachtet werden. Die Lagegenauigkeit gleichartiger Objekte oder Objektteile sollte möglichst homogen sein, damit signifikante analytische Auswertungen möglich sind.

Es wird geschätzt, dass etwa **80%** aller geschäftsrelevanten Daten einen Raumbezug haben! Dabei sind aber zu unterscheiden:

- Daten die Erscheinungen und Objekte des Geo-Raumes **geometrisch und attributiv** beschreiben, d.h. in ihrer **Lage, Form** und **Bedeutung**. Beispiele sind Flurstücke, Gebäude, Biotope, Naturschutzgebiete, Leitungen, ... → **Geodaten** *)

  *) ***Geodaten ≠ Grafikdaten: Geodaten sind mehr als Grafikdaten, Grafikdaten sind lediglich eine Präsentationsform von Geodaten, die man durch Zuweisung eines Styles erhält.***

  OGC: „We define a map as a visual representation of geodata, a map is not the data itself.“

- Daten über nicht geometrische Erscheinungen, die mit einem Ort (Koordinate) direkt oder indirekt verknüpft werden, zu ihrer Lokalisierung oder zum Erfassen ihrer räumlichen Verteilung (Karte mit „Fähnchen“). Beispiel sind Mieter, Eigentümer, Kunden, Verbraucher, Servicestellen, ... → **georeferenzierte Fachdaten**

Entwurf einer Taxonomie für Geoinformationssysteme

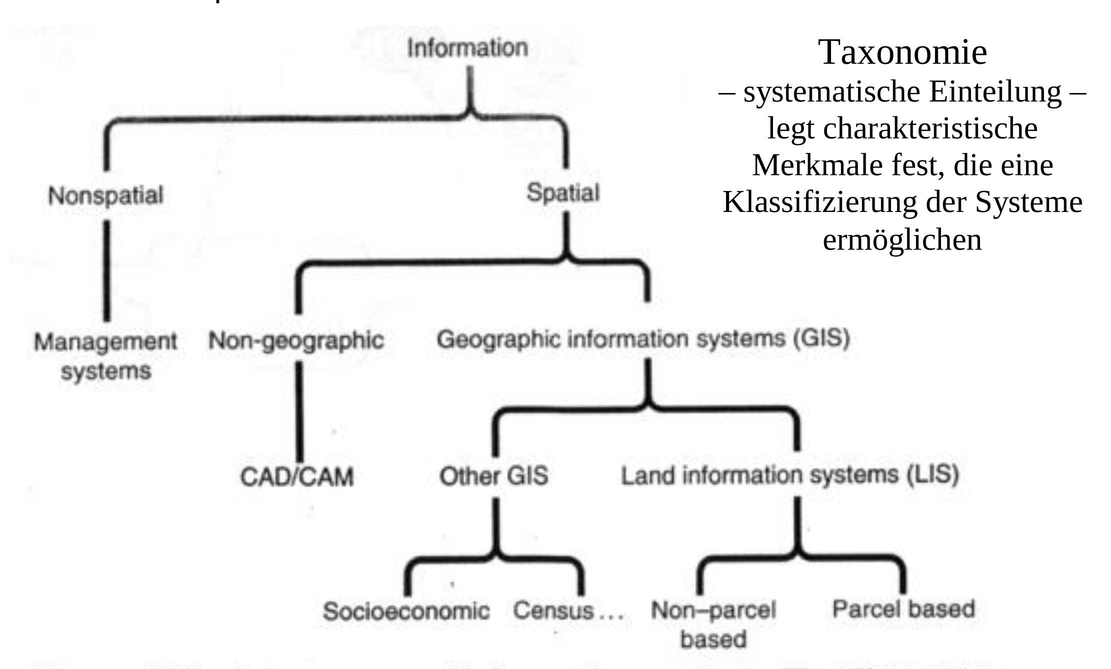
Im Folgenden werden generelle Merkmale von GIS aufgezeigt, die GIS-Einsätze aus verschiedenen Blickwinkeln beleuchten und ihre systematische Einteilung in „GIS-Arten“ erlauben sollen. Als Klassifizierungsmerkmale für GIS werden eingeführt:

- Fachanwendung
- Lebensdauer der Geodaten
- Räumlicher Bereich und Datenauflösung

} immanent

- Architekturtyp der IT-Komponenten → temporär

Daraus ergeben sich die Grundzüge einer Taxonomie, die helfen soll die Art eines GIS zu beschreiben und zu definieren und einem GIS einen konzeptionellen Rahmen mit arttypischen Schwerpunkten zuzuordnen.



Taxonomie
– systematische Einteilung –
legt charakteristische Merkmale fest, die eine Klassifizierung der Systeme ermöglichen

Figure 1.1 A taxonomy of information systems. The illustration shows how GIS and LIS fit in.

Fachanwendung

Für GIS ist die Einteilung nach ihrem Anwendungsgebiet am gebräuchlichsten. So gibt es heute eine ganze Reihe von GIS-Derivaten, wie z.B. kommunale GIS (KIS), Netzinformationssysteme (NIS), Umweltinformationssysteme (UIS). Dies sind GIS-Anwendungen in verschiedenen Fachbereichen.

Die Fachanwendung gibt die **Themen**, den Typ der **Geoobjekte,** ihre Merkmale und Beziehungen sowie die zugehörigen Qualitätsanforderungen vor, die in einem GIS verarbeitet und geführt werden sollen und die **Produkte**, die aus diesen Geodaten abgeleitet werden müssen. Wichtige konzeptionelle Unterschiede ergeben sich in Abhängigkeit vom Objekttyp. Für die Verarbeitung punkt-, linien- oder flächenhafter sowie künstlicher oder natürlicher Objekte sind hinsichtlich der Modellbildung und der benötigten Methoden unterschiedliche Anforderungen zu berücksichtigen. Für manche Anwendungen reichen Abbildungen der Objekte in Form von georeferenzierten Rasterdaten (Bilder) aus, während andere komplexe Modelle mit topologisch strukturierten Vektordaten in der Kombination mit physikalischen Parametern erfordern.

GIS kann man aus Sicht der Betreiber (Zuständigkeiten) grob in zwei Hauptsparten einteilen, in die Basisgeoinformationssysteme und in die Fachgeoinformationssysteme.

Basisgeoinformationssysteme

enthalten Geodaten, die für andere GIS-Anwendungen als Raumbezugssystem oder als georäumlicher Hintergrund benötigt werden, die

Geobasisdaten.

Ihre Führung obliegt i.d.R. den Vermessungsverwaltungen der Länder. Wichtige Quellen für Geobasisdaten sind das Amtliche Liegenschaftskataster-Informations-system (ALKIS) und das Topograpisch-Kartographische Informationssystem (ATKIS). Diese werden ergänzt durch landesweit verfügbare digitale Orthophotos (DOP) und digitale Geländemodelle (DGM) verschiedener Auflösung.

Fachgeoinformationssysteme

setzen mit ihren zusätzlichen Informationsebenen (Themen) auf diesen Geobasisdaten auf, deren Verfügbarkeit und Qualität ist ein wesentlicher konzeptioneller Aspekt beim Aufbau eines GIS.

Lebensdauer der Geodaten

Schwerpunkte von GIS-Anwendungen sind die raumbezogene Erfassung und Modellierung von Umweltressourcen für ihre Verwaltung und nachhaltige Nutzung sowie die von technischen Anlagen und Gütern als Grundlage für ihren Unterhalt, ihren Betrieb und ihre Planung. Der GIS-Einsatz kann dabei sowohl auf die langfristige Bestandsdokumentation als auch auf die Durchführung von einzelnen, unabhängigen Projekten ausgerichtet sein. Daraus resultieren hinsichtlich der Lebensdauer und der Modellierung der Geodaten unterschiedliche Systemkonzepte.

Bestandsdokumentation

GIS, die in erster Linie für die Bestandsdokumentation eingesetzt werden, enthalten meist Kernthemen für Anwendungsgebiete, die die Grundlage für viele Fragestellungen bilden. Die Bestandsdatenbank ist auf die langfristige Dokumentation der raumbezogenen Daten hin ausgelegt. Die Datenbestände müssen aus IT-Sicht über viele Hard- und Softwaregenerationen verwaltet und beherrscht werden und möglichst allen benötigten Verfahren und Methoden als Basis zur Verfügung stehen. Die Qualität der Datensammlung hat Vorrang; ihre Kontinuität muss über lange Zeiträume gewährleistet sein

- persistente Datensicht (dauerhaft, langfristig) -

Dazu ist ein vorausschauendes, zukunftsorientiertes Vorgehen erforderlich, denn ausschlaggebend sind Struktur, Informationsgehalt und Organisation der Datensammlung, sie bestimmen weitgehend das Einsatzspektrum und damit auch die Wirtschaftlichkeit eines Systems. Die eingesetzten Hard-/Software-Komponenten sind aufgrund der notwendigen, langen Lebensdauer (einige Jahrzehnte) der Geodatenbank nur von temporärer Bedeutung. Man kann davon ausgehen, dass die Hardware ca. alle 2-3 Jahre einen Leistungssprung nach oben macht und alle 5-7 Jahre vollständig ersetzt werden muss. Bestandsdatenbanken sollten wegen der Verfügbarkeit für ressort-übergreifende Mehrfachnutzungen, Redundanzvermeidung, Sicherheit und Sicherung der Daten möglichst zentral verwaltet werden.

Für den Einsatz der digitalen Datentechnik gilt allgemein, dass ihre Vorteile vor allem oder überhaupt erst dann eintreten, wenn größere Datenbestände aus langer Sicht und für eine möglichst vielseitige Verwendung angelegt werden, und zwar in Form von Informationssystemen.

Projektlösungen

Beim GIS-Einsatz für Projekte (z.B. Planung, Forschung) müssen die Prioritäten bei der Bewertung der Systemkomponenten anders festgelegt werden als bei der Bestandsdokumentation.

Ein Projekt ist nach *DIN 69901 „ein definiertes Vorhaben, detailliert geplant, mit genauer Finanzierung, das in einem festgesetzten Zeitrahmen abgewickelt wird“*, d.h.

Projekte sind abgeschlossene, zeitlich begrenzte Vorgänge.

Für ihre Durchführung müssen alle erforderlichen Geodaten bereitgestellt werden. Es wird eine fachübergreifende Benutzersicht aus den relevanten Datenbeständen benötigt. Für die Durchführung von Projekten werden z.B. spezielle Verfahren und Methoden für den Planentwurf, Simulationsmodelle, statistische oder betriebswirtschaftliche Auswertungen benötigt. Die verwendeten Werkzeuge (Verfahren, Methoden) müssen möglichst optimal geeignet sein, die projektspezifischen Aufgaben zu lösen. Die Durchführbarkeit der Aufgabe hat Vorrang. Das Vorgehen ist in Hinblick auf die effiziente Projektabwicklung weitgehend zweckorientiert. Die erzeugten Datenstrukturen sind nur von sekundärer Bedeutung solange sie ihren Zweck erfüllen.

- transiente Datensicht (zeitlich begrenzt) -

Projektlösungen können autonom eingesetzt werden; Ziel ist jedoch ihre Anbindung an vorhandene Geodatenbanken zur Bereitstellung und Übernahme von verfügbaren Datenbeständen. Die Erfassung und Bereitstellung der Geodaten verursachen oft erhebliche Kosten, sie können mehr als 80% der gesamten Kosten eines Vorhabens ausmachen.

Räumliche Ausdehnung und Datenauflösung

Wichtige Systemmerkmale für ein GIS sind seine räumliche Ausdehnung, d.h. **Dimension** und **Größe des Gebietes**, das erfasst wird und die **geometrische Auflösung** des Datenmodells (Abstraktionsgrad).

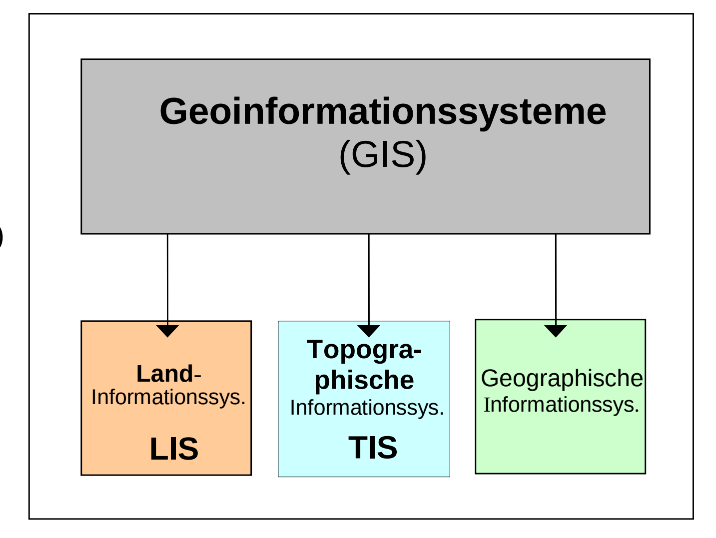
Unter Auflösung ist dabei der Grad der ”Feinheit” der Daten zu verstehen mit denen die Geoobjekte abgebildet werden. GIS können global, regional oder lokal mit zwei, drei oder vier Dimensionen aufgebaut werden.

In Abhängigkeit von ihrem Maßstabsbereich können GIS unterschieden werden in:

Landinformationssysteme (**LIS**) für den Inhalt eines Mehrzweckkatasters (Flurkarten, Stadtgrundkarten, technische Pläne für Leitungen, ...)
Maßstab: 1:1 - 1: 5.000
metrische Genauigkeit: cm (ALKIS)

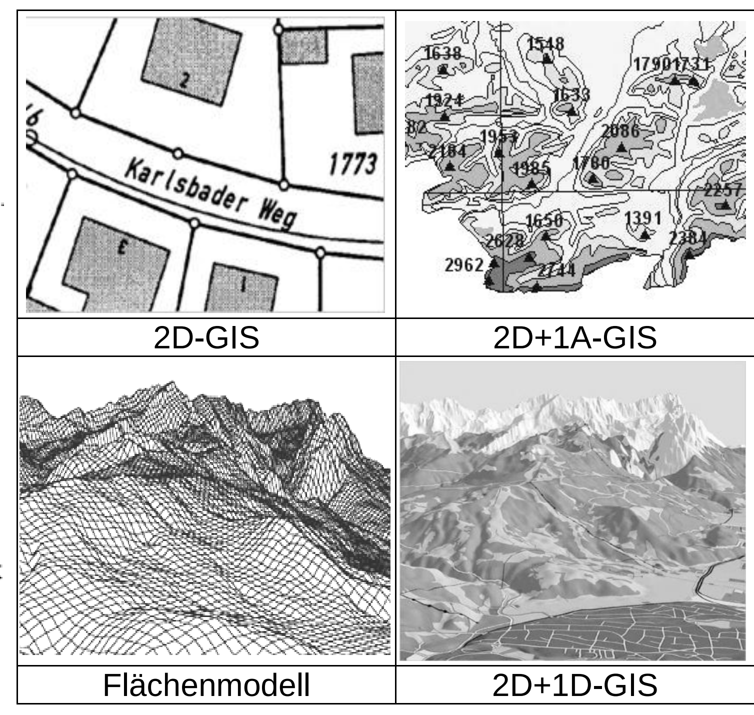
Topographische Informationssysteme (**TIS**) für die Abbildung der natürlichen und künstlichen Landschaft (Topographische Karten) Maßstab: 1:5.000 - 1:500.000, metrische Genauigkeit: m (ATKIS ± 3m)

Geographische Informationssysteme für geographische und thematische Inhalte (Geographische Karten)
Maßstab: 1:1.000.000 und kleiner,
metrische Genauigkeit: einige 100m



Geometrische Dimensionen

- **Zweidimensional (2D):**
  Geometriedaten lediglich x,y-Koordinaten (Planimetrie) ohne Höhenangaben
- **Zwei-plus-eindimensional (2D+1D):**
  Planimetrie + Digitales Geländemodell (DGM). I.d.R. ein Flächenmodell, das nicht mit der Lagegeometrie verknüpft ist
- **Zwei-plus-Attribut Höhe (2D+1A)**
  Speicherung der Höhe z als Attribut der Lagegeometrie.
  (In Bereichen ohne Lagegeometrie erfolgt keine Aussage)
- **Dreidimensional (3D):**
  Speicherung der x,y,z-Koordinaten in hinreichender Dichte.
  Weitere Unterscheidung in Linienmodell, Flächenmodell und Volumenmodell
- **Vierdimensional (4D):**
  Zusätzliche Speicherung des Zeitparameters t neben den x,y,z-Koordinaten.
  Anwendung: Zeitliche (kinematische) Dokumentation geowissenschaft. Fragen



2D-GIS | 2D+1A-GIS | Flächenmodell | 2D+1D-GIS

2D+1A = Lage mit Höhenwert als Attribut

2D+1D = Grundriss und parallel dazu ein Flächenmodell zur Interpolation der Höhe

Räumliche Auflösung *(Level of Detail =LOD)*

Die räumliche (geometrische) Auflösung bezeichnet die „Feinheit“ oder „Schärfe“ mit denen Objektstrukturen erfasst sind. Während dieser Begriff bei Rasterdaten gebräuchlich ist, muss für Vektordaten ein Äquivalent eingeführt werden. Die räumliche Auflösung von Vektordaten kann durch die Datendichte und der metrischen Genauigkeit der Koordinaten beschrieben werden, wobei beide Einflussfaktoren weitgehend vom Erfassungsmaßstab abhängen.

Wichtige Maßstabsbereiche für Detail- und Rahmenpläne sowie Übersichten sind:

• 1: 1.000	Standardmaßstab in bebauten Gebieten,	Kataster
• 1: 5.000	Grenzmaßstab für Zusammenführung von LIS und TIS	**Detailpläne**
• 1: 10.000	Grenzmaßstab für die grundrisstreue Darstellung	
• 1: 25.000	schwach generalisierte, grundrissähnliche Darstellung	Topographie
• 1: 50.000	generalisierte Darstellung, z.B. aus der Landes- und Regionalplanung	**Rahmenpläne**
• 1:100 000	stark generalisierte Darstellung, Grenzmaßstab zu geographischen Anwendungen.	Atlas **Übersichten**

Die Größe eines Einsatzgebietes in Verbindung mit der erforderlichen Auflösung der Geodaten ergibt wichtige Hinweise über Zeit-, Kosten- und Ressourcenbedarf (z.B. Speicherplatz für Datenbanken) beim Aufbau eines GIS.

Es ist bei der Konzeption zu beachten, dass die für die Abbildung eines Sachverhalts nötige Datendichte sich annähernd quadratisch zur Änderung des Maßstabs verhält, die Wahl des richtigen Erfassungsmaßstabs hat damit wesentlichen Einfluss auf den Aufwand (Zeit und Kosten) für die Datenerfassung.

IT- Komponenten

Eine weitere, gebräuchliche Einteilungsmöglichkeit bietet die Gruppierung von GIS nach den eingesetzten IT-Komponenten (Systemarchitektur), so können je nach Größe und Aufgabe des Systems z.B.

- unternehmensweite, verteilte Client-Server-Architekturen
- Workstation für komplexe Analysefunktionen
- Desktop und Web-Technik für die breite Anwendung im Auskunftsbereich

zum Einsatz kommen.

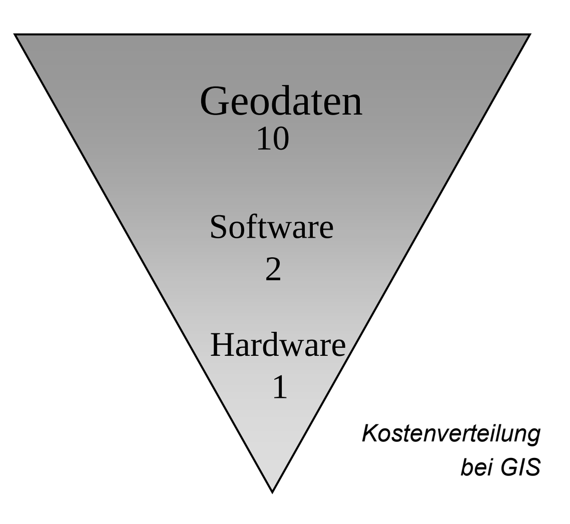
Während die bereits aufgeführten drei Klassifizierungskriterien unveränderliche (immanente) Merkmale der Geodaten für eine Anwendung sind, ist die Auswahl der eingesetzten Systemarchitektur ein temporaler Aspekt, der neben dem Benutzprofil stark von der zur Zeit der Einführung verfügbaren IT-Technologie abhängt.

Bei GIS, die langfristig angelegt werden, ist innerhalb ihrer Lebensdauer i.d.R. mit mehreren Hard- und Softwaregenerationen zu rechnen. Art und Verteilung der eingesetzten IT-Komponenten sollte beim Ersatz einer Generation (ca. alle 5 Jahre) in Abhängigkeit vom wachsenden Aufgabenspektrum und den verfügbaren Kommunikationsmöglichkeiten an den Stand der Technik angepasst werden. Bei großen GIS-Applikationen kommen meist Kombinationen verschiedener Systemarchitekturen zum Einsatz. Der Aufwand für Migrations- und Portierungsarbeiten bei einem Wechsel der IT-Komponenten ist ein wichtiger Faktor für die Wirtschaftlichkeit und Lebensdauer eines GIS.

Anwender binden sich bei Systemen für die Bestandsdokumentation wegen der teuren Einführungs-, Entwicklungs- und Einarbeitungsphasen sowie der kostspieligen Datenerfassung und den meist proprietären Datenmodellen meist längerfristig an eine GIS-Software. Die richtige Produktwahl und Produktkombination hat deshalb einen wesentlichen Einfluss auf den erfolgreichen Verlauf des Systemseinsatzes über längere Zeiträume. Besonders wichtig bei der Produktauswahl sind daher auch Fragen wie die Marktpräsenz des Herstellers und die Einsatzschwerpunkte der ausgewählten GIS-Software.

Das weite Feld der Erfassung, Verwaltung und Präsentation von Geodaten kann i.d.R. nicht mit einem GIS-Produkt optimal abgedeckt werden.

Aus der Relation der Kosten ist ersichtlich, dass die Hardwarekomponenten in der Gesamtbilanz nur ein geringes Sparpotential beinhalten. Eine falsche Sparsamkeit bei der Auswahl der Hardware ist deshalb nicht ratsam. Die Zuverlässigkeit und Verfügbarkeit, aber auch die Ergonomie der Hardware sind wichtige Produktionsfaktoren. Ausfallzeiten oder die Leistungsabnahme beim Personal, wegen ergonomisch ungünstiger Geräte, können die eingesparten Mittel bei der Beschaffung von „Billigprodukten“ schnell um ein Vielfaches durch erhöhte Betriebskosten übersteigen.



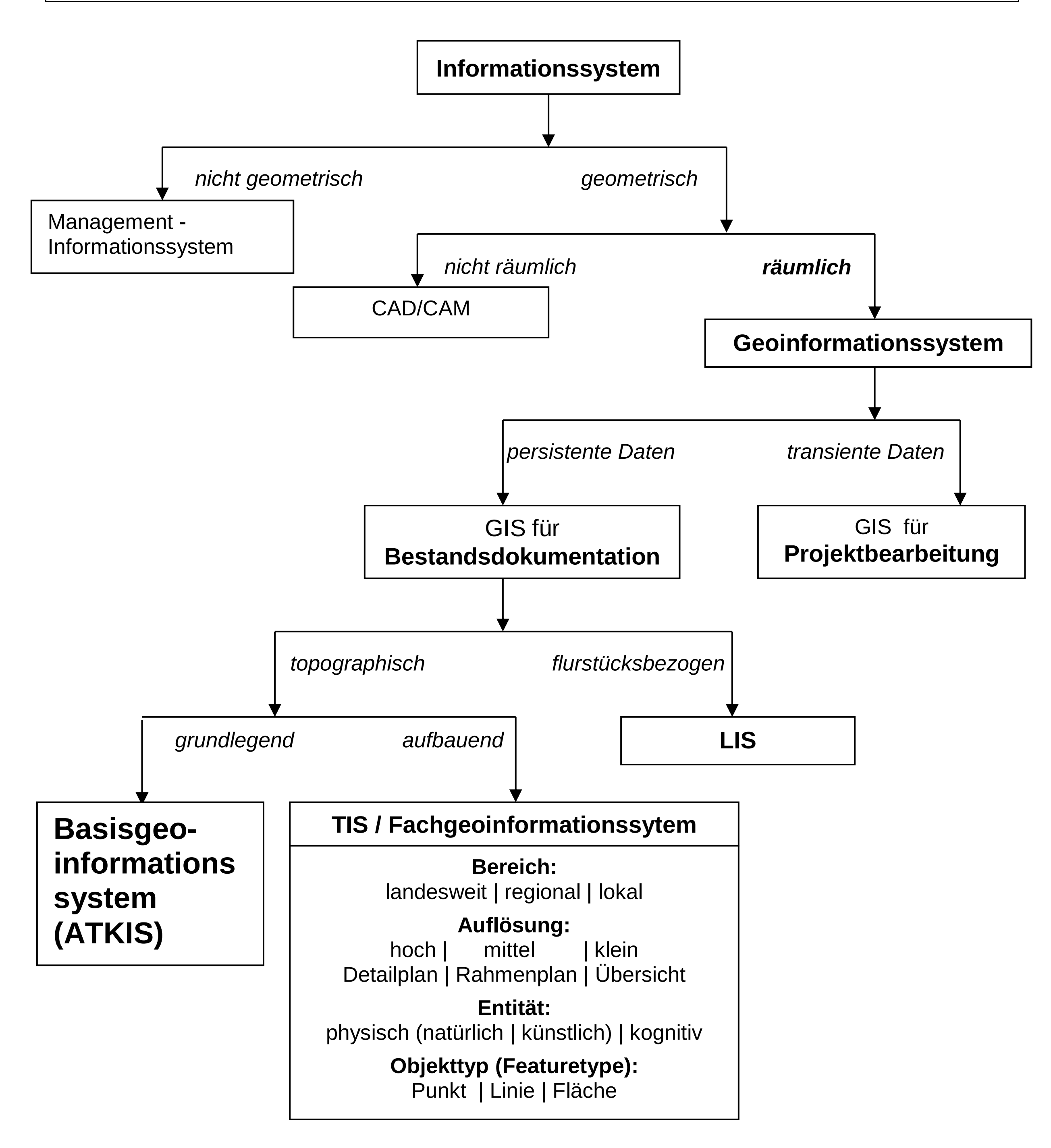
Kostenverteilung bei GIS

Beispiel für die Klassifikation eine GIS (Fach-GIS)

- **Fachanwendung**
- **Lebensdauer der Geodaten**
- **Räumliche Bereich und Datenauflösung**
- **Architekturtyp der IT- Komponenten**

Klassifizierung ergibt wichtige Rahmenbedingungen für ein Grobkonzept eines GIS

IT-Architektur =
f [(Fachanwendung, Lebensdauer, Bereich/Auflösung), verfügbare IT-Technik]



FORST-GIS = Topographisches Fachgeoinformationssystem für die landesweite Bestandsdokumentation forstlicher Flächeninformation (mit verteilten Systemkomponenten)

Abgrenzung von GIS gegenüber anderen Systemen

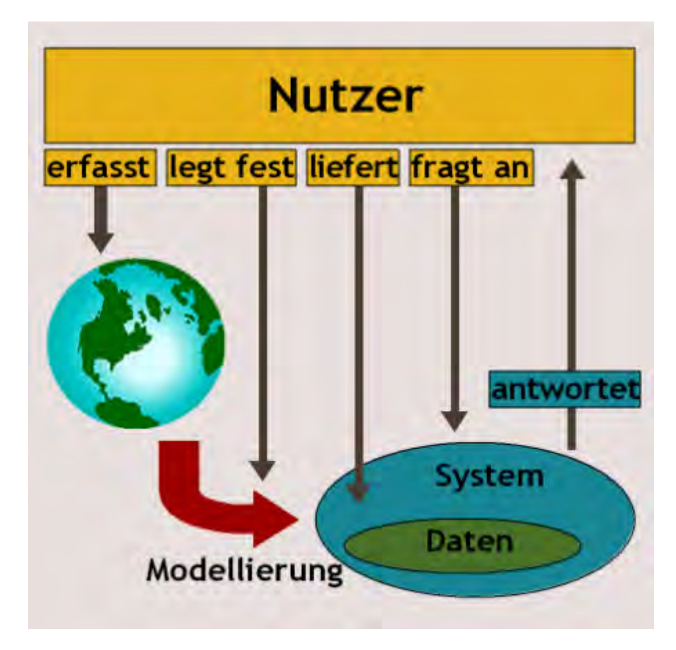
GIS versus IS (DB)	
Geoinformationssystem	Informationssystem (Datenbank)
• Geoobjekte mit explizitem Raumbezug und gekoppelten Sachdaten • Selektion von Geoobjekten über Raumbezug und Attribute • Datenanalyse interaktiv-grafisch, numerisch-statistisch • Visualisierung mit digitaler Kartographie, Tabellen und Diagrammen	• Modellierung allgemeiner Objekte, Raumbezug nur als Attribut • Selektion von Objekten nur über Attribute möglich • Datenanalyse überwiegend mit statistischen Methoden • Visualisierung durch Tabellen und Diagramme (Business-Grafik)
GIS versus CAD	
Geoinformationssystem	Computer Aided Design
• Abbildung der Realität durch ein geometrisch und fachlich vereinfachtes Modell • Geometrie und Thematik der Geoobjekte sind gekoppelt • Geoobjekte sowohl in Vektor- wie im Raster-Modell darstellbar • Analysefunktionen bilden den Schwerpunkt der GIS-Funktionalität	• „Von der Idee zur Realität" durch interaktiv-geometrisches Modellieren und Konstruieren • Meist keine Sachdaten-Verwaltung erforderlich • Geometrie der Objekte nur vektoriell sinnvoll • Analysefunktionen in der Regel nur rudimentär vorhanden
GIS Modellierung natürlicher und technischer Erscheinungen (Entitäten) des Geo-Raumes in Form von Geoobjekten durch die approximative Abbildung ihrer geometrischen Form und Beschreibung ihrer Merkmale mittels Attributen	CAD exakte Abbildung der Geometrie künstlicher Objekte im Rechner mit dem Ziel der Produktionssteuerung (CAD/CAM); Fertigungsroboter
GIS versus Kartographiesystem	
Geoinformationssystem	Kartographiesystem (Mapping System)
• Modellierung der realen Welt • Visualisierung mit grafischen und kartographischen Techniken ist nur die Präsentationsfunktion • Selektions- und Analysefunktionen bilden den Schwerpunkt der GIS-Funktionalität	• Modellierung der Karte (2D-Abbild) • Primäres Ziel ist die Konstruktion topographischer und thematischer Karten • Analysefunktionen sind nur eingeschränkt vorhanden

Heute ist eine immer stärkere Konvergenz und funktionale Integration dieser konzeptionell unterschiedlichen Systemtypen zu beobachten, da sie im Wesentlichen auf dieselbe IT-Technik aufbauen. → Dieselbe IT-Technik wird für die Bearbeitung unterschiedlicher Fachaufgaben, mit unterschiedlichen Daten eingesetzt.

3 Datenstrukturen für Geodatenmodelle

Die Geodaten sind der Kern eines GIS. Die Fragen der Datenmodellierung bilden daher einen Schwerpunkt der Vorlesung Geoinformatik. Das optimale, bedarfsgerechte Design der Geodaten ist die grundlegende Voraussetzung für ein hohes Informationspotential und damit für ein breites Einsatzspektrum der Geodatenbank. Versäumnisse bei der Datenmodellierung können später oft nur mit großem Aufwand ausgeglichen werden, deshalb ist ein sorgfältiger, bedarfsgerechter Entwurf der Datenstruktur eine wichtige Voraussetzung für GIS, besonders bei langfristig angelegten Bestandsdatenbanken.

Datenmodellierung bezeichnet die Abbildung der Realität auf Strukturen und Prozesse eines Informationssystems.



Das Modell legt fest, was das Informationssystem "weiß"

Ein Modell begründet einen "Vertrag" zwischen Nutzer und Informationssystem:

- Welche Daten muss der Nutzer liefern?
- Welche Anfragen will der Nutzer stellen?
- Welche Antworten soll das System geben?



Die vier Sichten (Entwürfe) von Geodatenmodellen



- **Externes Schema:** Die fachliche Benutzersicht (Teilsichten, View) ergibt sich aus der Inhaltsanalyse der relevanten Fachdaten.



- **Konzeptionelles** (konzeptuelles) **Schema:** Die konzeptuelle Gesamtsicht, sie entsteht aus der Synthese der fachlichen Teilsichten. Es ist eine formalisierte Beschreibung der Welt, **das semantische Datenmodell**.
  (beide Bezeichnungen üblich: **konzeptionell** → ein Konzept betreffend, **konzeptuell** → ein Konzept aufweisend)
- **Logisches Schema:** Der logische Entwurf erfolgt auf Basis der in den eingesetzten GIS- und Datenbankprodukten verfügbaren grundlegenden Datenstrukturen (z.B. GDM, RDM, OODM), mit denen das semantische Datenmodell adaptiert wird.
- **Internes Schema:** Die physische Sicht beschreibt die Implementierung und Generierung der Geodatenbanken auf Ebene der Hardware, des Betriebssystems und der Datenbankprodukte (Dateien, Datenblöcke, Partitionierung, Fragmentierung, Zugriffsmechanismen, Indextabellen, ...).

Bei der Konzeption und dem Entwurf von Geodatenmodellen werden zunehmend vier statt drei Datenbankschemata genannt. Das konzeptionelle Schema wird weiter aufgeteilt in das eigentliche, abstrakte konzeptionelle Schema, das von der Implementierung unabhängig ist, und in das logische Schema zur konkreten Modellierung der Geodaten, das von der eingesetzten Software abhängig ist.

Die Aufteilung der Gesamtsicht einer Datenbank in das konzeptionelle und das logische Schema hat den Vorteil, dass der konzeptionelle Entwurf in einer von der Realisierung unabhängigen Form erfolgen kann. Heute werden dafür meist objektorientierte Notationen wie die Unified Modeling Language (UML) oder das Entity-Relationship-Modell (ERM) verwendet.

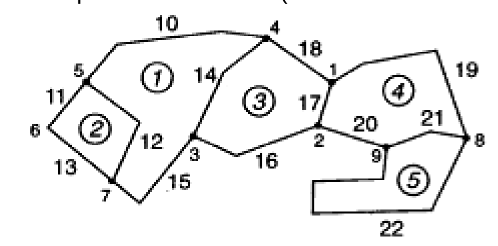
Das Ergebnis des konzeptionellen Entwurfs ist das konzeptionelle Schema, das auch als **semantisches Datenmodell** bezeichnet wird.

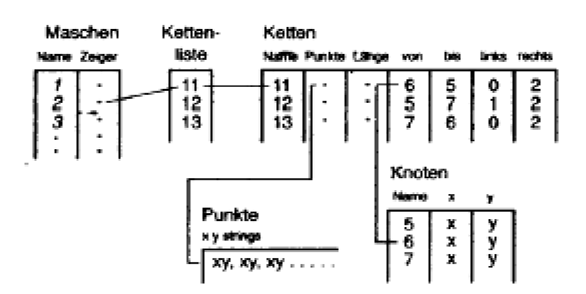
Auf Basis des konzeptionellen Schemas wird der logische Entwurf durchgeführt, dessen Ergebnis mit der im ausgewählten Datenbanksystem verfügbaren Datendefinitionssprache (DDL) dargestellt werden kann.

Wie die Objekte und ihre Beziehungen aus dem konzeptionellen Schema in das logische Schema überführt werden können, hängt davon ab, welche Datentypen und Datenstrukturen die eingesetzte GIS-Software und die Datenbanksoftware unterstützen.

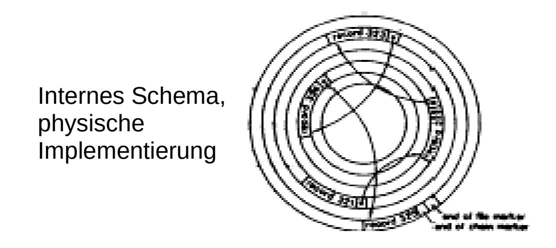
Modellbildung

Konzeptionelles Schema (semantisches Modell)





↑ Logisches Schema (z.B. relationales Modell)



Internes Schema, physische Implementierung

Die logische Modellierung von Objekten, insbesondere die von Geoobjekten, ist nicht standardisiert, proprietäre Lösungen sind deshalb heute die Regel → keine Interoperabilität.

Datentypen für Geodaten

Die Modellierung des Geo-Raumes in einem GIS erfordert komplexe, vernetzte Datenstrukturen, d.h. unterschiedliche Datentypen, um die geometrischen und fachlichen Merkmale der Geoobjekte sowie die Beziehungen zwischen Geoobjekten entsprechend den fachlichen Anforderungen optimal zu beschreiben.
Folgende Datentypen sind dazu erforderlich:

- **Vektordaten** für Geometrie (Geometrieelemente: Punkt, Linie, Fläche) und die kartographische Ausgestaltung (Text, Symbol)
- **Topologische Verweise** (Knoten, Kante, Masche) für Nachbarschaftsbeziehungen
- **Attribute** und Texte zur Beschreibung physikalischer, betriebswirtschaftlicher, ökologischer und soziologischer Parameter (Tupel in Tabellen)
- **Rasterdaten** für Bildinformation (Pixel, Bildmatrix) oder andere flächenhaft in Rasterzellen (Grid) vorliegende Information
- **Digitale Höhenmodelle** (DHM) [Geländemodelle (DGM), Oberflächenmodelle(DOM)]

Für den gemeinsamen Einsatz unterschiedlicher Datentypen, insbesondere von Vektor- und Rasterdaten, wird auch der Ausdruck **hybride Datenverarbeitung** verwendet. Die Attribute werden häufig als **Sachdaten** bezeichnet.

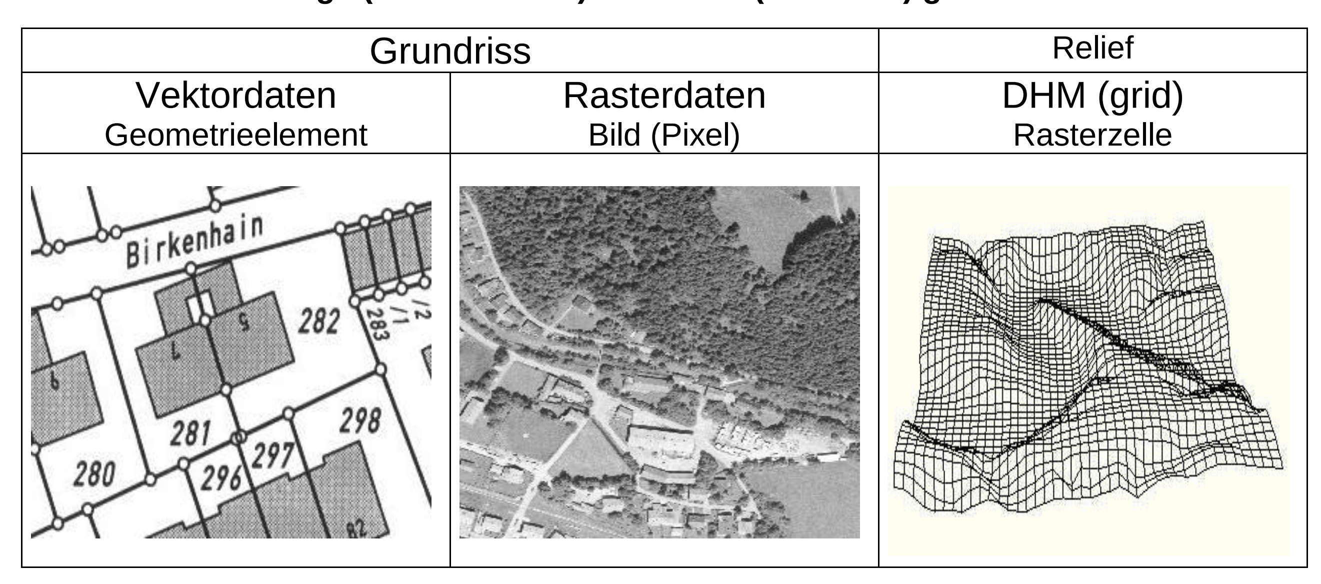
Geometriemodelle

Für die Beschreibung geometrischer Merkmale werden in GIS meist zweidimensionale Modelle verwendet, da die Hauptanwendung von GIS die raumbezogene Beschreibung von ausgedehnten Entitäten der Eroberfläche ist, die zweidimensional modelliert werden können. Die direkte dreidimensionale Modellierung im GIS ist heute noch ein Spezialfall, der sich auf Anwendungen z.B. in der Stadtplanung oder Geologie beschränkt. Höhenangaben haben in GIS gegenüber den Lageangaben als geometrische Dimension meist eine nachgeordnete Bedeutung, sie können ohne Informationsverlust als Ergänzung zu den zweidimensionalen Strukturen in Form von Attributen an Punkten (2½D) oder als zusätzliches DGM (2D+1D) modelliert werden.

Die geometrische Modellierung im GIS wird hier für den zweidimensionalen Regelfall dargestellt und bei Bedarf, im Rahmen der Beschreibung von GIS-Anwendungen, auf den Spezialfall mehrdimensionaler Modelle verallgemeinert.

Die Abbildung der meisten raumbezogenen Sachverhalte im GIS erfolgt nach Lage (2D-Grundriss) und Höhe (1D-Relief) getrennt

Grundriss		Relief
Vektordaten Geometrieelement	Rasterdaten Bild (Pixel)	DHM (grid) Rasterzelle



Die Abbildung der geometrischen Information kann in GIS mit Vektor- oder Rasterdaten erfolgen. Welcher Datentyp verwendet wird hängt davon ab, welche Eigenschaften und Merkmale einer Entität modelliert werden sollen und welche Datenerfassungsmethoden zur Verfügung stehen. Für flächenhafte kontinuierliche Information ist das Rastermodell, für linien- oder flächenhafte diskrete Information das Vektormodell zu bevorzugen.

Vektormodelle („lineale Modelle“)

Das 2D-Vektormodell beinhaltet Punkte, linienhafte Verbindungen dieser Punkte und Flächen, die von Linienkonturen gebildet werden (Randbeschreibung). Flächen können nur in Verbindung mit Linien, Linien nur in Verbindung mit Punkten existieren.

Das geometrische Grundelement der Vektordaten ist der Punkt, der durch die Angabe seiner Koordinate (Ortsvektor) eindeutig definiert ist.

Die Beschreibung von Linien erfolgt durch die Approximation mittels Stützpunkten, die geradlinig (Polygonzug, Polyline) oder mittels interpolierender Kurven (Spline) verbunden werden. Flächenstrukturen (Polygone) werden durch ihre Begrenzungslinien (Kontur) beschrieben. Die metrische Information von Vektormodellen steckt in den Koordinaten der Punkte bzw. den Stützpunkten der Linien.

Als physische Datenstruktur für Vektoren werden Listen verwendet. Vektoren können

- einfach strukturiert sein – als sogenannte Spaghettis (lineare Liste), dies sind Linien die durch einfache (lange, dünne) Koordinatenlisten beschrieben werden
- topologisch strukturiert sein (z.B. mittels Ringlisten), mit einer Knoten-Kanten-Struktur, entsprechend ihrer Inzidenzbeziehungen

Für Geodatenbanken ist wegen der erforderlichen Datenkonsistenz die Verwendung topologisch strukturierter Vektordaten zu empfehlen (siehe Kapitel Geoobjekte).

Vor-/Nachteile von Vektordaten:

- **Diskretisierbare Entitäten mit scharfen Konturen** lassen sich gut mit Vektordaten modellieren, sie sind jedoch weniger gut geeignet für flächenhafte Erscheinungen mit unscharfen Übergängen.
- Vektordaten sind gut für die Randbeschreibung von Geoobjekten geeignet.
- Geoobjekte sind mit Vektordaten praktisch mit beliebig hoher geometrischer Genauigkeit der Lage und Form nach darstellbar, allerdings wird der Erfassungsaufwand für die Diskretisierung bei zunehmender Genauigkeit sehr hoch.
- Logische und algebraische Operationen mit Vektordaten sind numerisch aufwendig, Koordinatentransformationen lassen sich dagegen einfach berechnen.
- Die Datenmengen von Vektormodellen sind im Vergleich zu Rastermodellen meist wesentlich geringer.

Rastermodell („areale Modelle“)

Die Eigenschaften von flächenhaften Entitäten können in GIS durch regelmäßige, in Rasterform angeordnete Bereiche (Zellen) beschrieben werden. Innerhalb einer Zelle werden die Eigenschaften als homogen betrachtet und durch einen einzigen diskreten oder skalaren Wert oder eine Funktion beschrieben.

Rasterdaten bilden ein Mosaik, das einer geometrisch verifizierbaren Abbildung der Erdoberfläche entsprechen muss. Meist werden quadratische oder rechteckige Maschen verwendet, die man als Rasterzelle (grid) oder im Falle von Bilddaten als Pixel (picture element) bezeichnet. Die physische Datenstruktur von Rasterdaten ist eine Matrix, sie werden deshalb auch als Array-Daten (Pixel, Grid) bezeichnet.

Flächenhafte Phänomene werden häufig zunächst durch unregelmäßige, diskrete Stützstellen (Messungen) repräsentiert, die dann mittels Interpolation auf regelmäßige Bezugsräume (grid) bezogen werden. Eine andere Möglichkeit zur Modellierung flächenhafter Strukturen sind Triangulationen mit Dreiecksmaschen (siehe DGM, TIN).

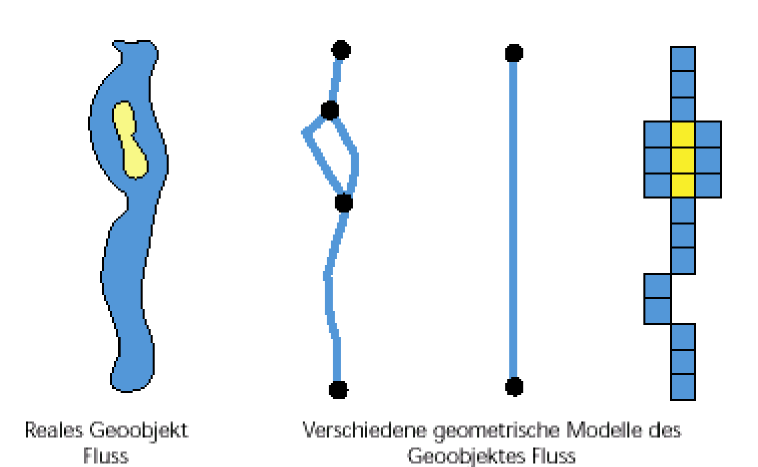
Vor-/Nachteile von Rasterdaten:

- **Flächenförmige Entitäten (Kontinua) mit hoher Variation oder unscharfen Übergängen** (z.B. Oberflächen, Vegetation) lassen sich gut mit Rasterdaten erfassen
- Die Rastergröße muss den zu erfassenden minimalen Strukturen angepasst werden.
- Diskrete Entitäten können bezüglich ihrer Geometrie nur approximativ (weder lagegenau noch formtreu) dargestellt werden, der Datenerfassungsaufwand z.B. mittels Scannern ist allerdings gering
- Logische und algebraische Operationen sind auf Basis von (koinzidenten) Rastern sehr einfach, Koordinatentransformationen sind aufwändig
- Die anfallenden Datenmengen sind im Vergleich zum Vektormodell sehr groß

Datenstrukturen („Bausteine“) für 2D-Geodatenmodelle

Einige generelle Anforderungen:

- Geodatenmodelle beinhalten primär geometrische Strukturen, die durch Attribute in ihrer fachlichen Bedeutung beschrieben werden. Die Basis bildetet ein Raumbezugssystem, mit einer definierten mathematischen Grundlage.
- Aus wirtschaftlicher Sicht, insbesondere wegen des Umfangs der Datenerfassung, ist darauf zu achten, dass die verwendete Datenstruktur einerseits komplex genug ist für die Wiedergabe der tatsächlich existierenden Beziehungen in dem zu modellierenden Sachverhalt, andererseits einfach genug für einen effizienten und performanten Betrieb des Systems ist.
- Das Modell muss so beschaffen sein, dass sein Informationsgehalt mit dem Basiswissen einer bestimmten Ziel-Benutzergruppe abgefragt werden kann und relevante Informationen für diesen Benutzerkreis mit einem vertretbaren Aufwand gewonnen werden können.


Reales Geoobjekt Fluss

Verschiedene geometrische Modelle des Geoobjektes Fluss

Basisdatenstrukturen:
- Geometrieelemente
- Rastermatrix
- Attribut,Tupel

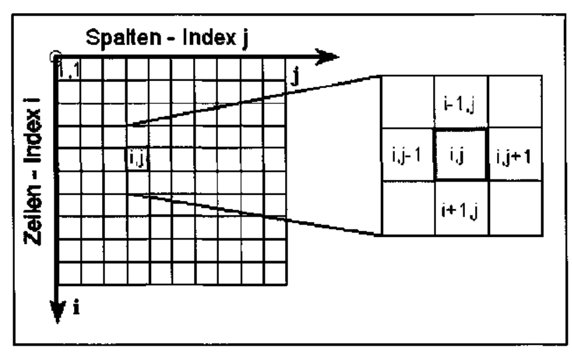
sind die elementaren Strukturen zur Abbildung der Eigenschaften von räumlichen Entitäten im GIS.

Geometrieelemente (Spatial Feature) sind die grundlegenden Datenstrukturen, für die approximative geometrische Modellierung von raumbezogenen Entitäten.

Bei einem 2D-GIS werden für die Grundrissdarstellung von Geoobjekten die Geometrieelemente Punkt, Linie, Fläche zur Beschreibung der Form und Lage sowie Text und Symbol zur kartographischen Ausgestaltung vorausgesetzt (Vektordaten).

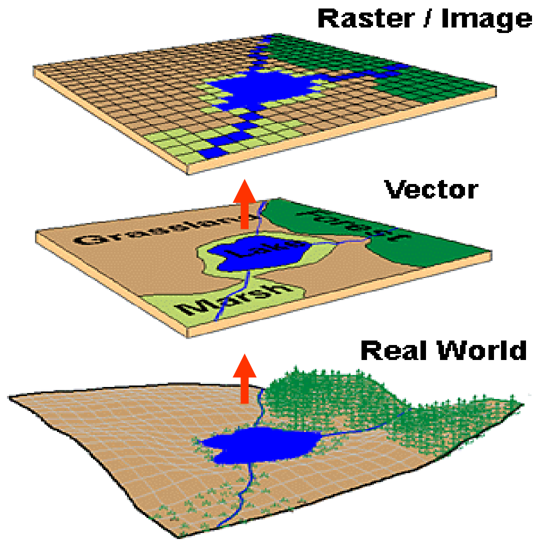
Elementtyp (Featuretype)	**Dimension**	Geometrie (spatial Feature)	Topologie	Verwendung
Punkt (Point)	0	Position, Koordinate	**Knoten** (Node)	Raumbezug, Punktobjekte
Linie (Line, Curve)	1	Gerade Polygonzug (Polyline,Arc) Spline, Kreisbogen, ...	**Kante** (Edge)	Kontur, Grenze Linienobjekte
Fläche (Area, Polygon)	2	Polygon (geschlossen)	**Masche** (Face)	Bezugsraum, Flächenobjekte
Text (Annotation)	0	Koordinate mit Text	singulärer Knoten	Ausgestaltung
Symbol	0	Koordinate mit Grafik	singulärer Knoten	Ausgestaltung

Rasterdaten sind die grundlegende Datenstruktur für die Modellierung ausgedehnter flächenhafter Phänomene, sie werden als Matrix (Array) verwaltet.



Die Rastermatrix bezeichnet die in Zeilen und Spalten regelmäßig angeordneten Rasterzellen. Es werden meist quadratische Rasterzellen verwendet, die bei Bildern das kleinste Bildelement (Pixel) darstellen. Zur Festlegung der Georeferenz einer Rastermatrix müssen ihr Ursprung, ihre Orientierung und die Zellgröße festgelegt werden.

Jeder Rasterzelle wird - als unteilbares, flächenhaftes Basiselement - ein Wert zugeordnet. Dem Raster-Modell liegt keine kontinuierliche Metrik wie den Vektoren sondern eine diskrete Metrik zugrunde. Zur geometrischen Lagebeschreibung einer Zelle kann man daher an Stelle der Koordinaten-Tupel (x,y) die Index-Tupel (i,j) verwenden, sie beschreiben die Lage einer Rasterzelle eindeutig. Als Ursprung wird meist die linke obere Ecke der Matrix verwendet.



Attribute sind die grundlegende Datenstruktur für die Modellierung der semantischen Merkmale und Beziehungen von Entitäten (Sachdaten). Attribute haben einen Namen, einen Datentyp und einen zugeordneten Wertebereich (Domäne).

Räumliche Entitäten können nicht vollständig geometrisch beschrieben werden. Für die Modellierung ihrer fachlichen Eigenschaften und sonstigen thematischen Merkmale, den sogenannten Sachdaten, sind zusätzlich Beschreibungen mittels Attribute erforderlich. Je nach Sachverhalt sind dazu unterschiedliche Attributtypen für qualitative oder quantitative Merkmalswerte erforderlich.

Attributtyp	Datentyp	Bemerkung
nominal	Text, Logical	lexikalisch, bezeichnen
ordinal	Text, Zahl	Relationen größer, kleiner
kardinal	Zahl	Differenz definiert
metrisch	Dezimalzahl	Strecke, Winkel
komplex	Medium	Bild, Ton, Video, Web-Link

Die Datentypen können Texte, Zahlen oder auch Bilder sein, abhängig davon, ob die Beschreibung einer Eigenschaften durch nominale, ordinale, kardinale oder metrischer Werte erfolgt. Die Semantik eines Attributs ergibt sich aus seinem Namen und seinem Wert sowie eventuell aus Zusatztabellen, falls an Stelle der tatsächlichen Attributwerte Schlüsselzahlen verwendet werden.

Die Attribute können zu Tupeln (Sätze) aggregiert werden, dies ist eine Zusammenfassung einer festen Anzahl benannter Attribute zu einem Datensatz mit den semantischen Merkmalen einer Entität.

Thema

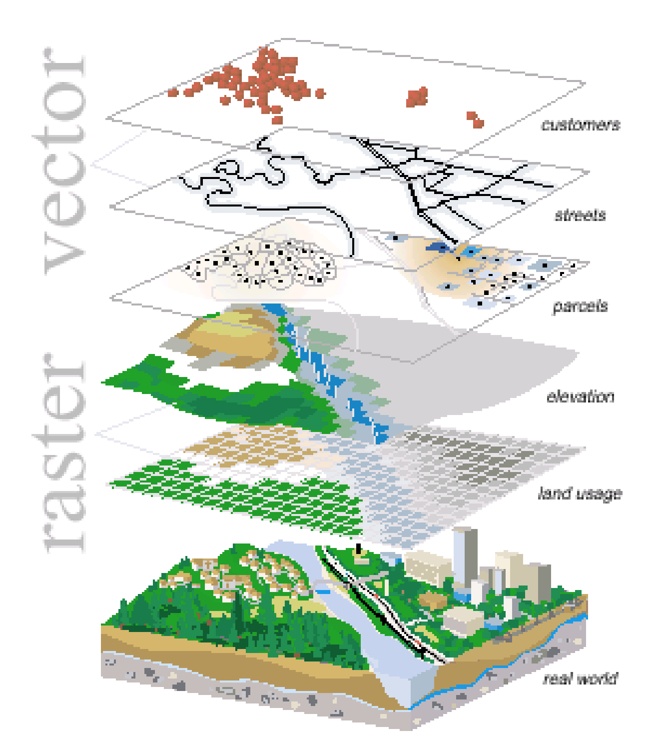
(unabhängige inhaltliche Schicht)

Ein Thema ist eine fachlich unabhängige Teilsicht des Gesamtmodells. Es ist ein Datensegment, dessen Inhalte logisch und organisatorisch zusammengehören.

Verschiedene Themen stehen nur über ihren gemeinsamen Raumbezug miteinander in Verbindung, sie haben a priori keine anderen Beziehungen.

Die geometrische Information eines Themas kann als **V**ektor- oder **R**asterdaten abgebildet werden, deren fachliche Bedeutung durch **A**ttribute beschrieben wird.

Jedem Thema ist ein Raumbezugssystem und ein Abstraktionsgrad sowie Verfahren für die Erfassung, Fortführung und Auswertung der Geodaten zugeordnet.



Beispiel FORST-GIS: Thematische Gliederung					
Ressort	**Thema**	Struktur	Typ	Verw.	Format
Forsteinrichtung	Forstbetriebskarte	F/L/**O**	V/M10	DB, p	SQD
	FE/DB-Begang	A	T	DB, t	SQD
Standortserkundung	Standortskartierung	F/L/**O**	V/M10	DB, p	SQD
	Hanglabilitätskarte	F	V/M25	D	ARC
Inventur	FE/DB-Inventur	A	T	DB, t	SQD
	Inventurgitter	F/**O**	V	DB, t	SQD
Schutzwaldsanierung		F/L	V/M10	DB, p	SQD
Forstorganisation	Forstorganisation	F/L	V/M50	DB, p	SQD
	Waldbesitzarten	F/L	V/M50	DB, p	SQD
Naturschutz	Waldfunktionen	F/L	V/M50	DB, p	SQD
	Amtl. Schutzgeb.	F/L	V/M50	DB, p/n	SQD
	FFH/SPA-Gebiete	F/L	V/M50	DB, p/n	SQD
Geobasisdaten	DTK25	L, nur für Projekte	R/M25	D	TIFF/B
	DTK50	L	R/M50	D	GTIFF/B
	DTK200	L	R/M200	D	GTIFF
	DGM, FK5-Höhe	L, nur für Projekte	R/M10	D	TIFF/B
	DOP	L, nur für Projekte	R/M5	D	GTIFF
	DFK	F/L	V/M5	DB, t	SQD

Struktur F = Geometrieelement / L = Layer / **O** = fachliches Objekt / A = Attribut
Format V = Vektor / R = Raster / T = Tabelle Mxx = Maßstabszahl (xx =Tausender)
Status p = persistent in DB / t = transient in DB / D = Datei / n = Nachrichtlich

Ein Thema kann in einer eigenen Geodatenbank oder in Dateien verwaltet werden. Themen bilden die Grobstruktur der Geodaten in einem GIS. Sie werden aus fachlichen und organisatorischen Gesichtspunkten heraus gebildet.

Beispiel: ESRI-Shapedaten = Themen

Bei umfangreichen Geodatenbeständen sind zwischen der Grobstruktur Thema und den Basisdatenstrukturen noch weitere Möglichkeiten zur Bildung von Organisationseinheiten gebräuchlich. Wichtige Konzepte sind die Layerstrukturierung der Geometrie und die objektstrukturierte Abbildung fachlicher Entitäten sowie ihre Kombination.

Layer (verknüpfte inhaltliche Schicht) - **vertikale Struktur von Themen**

Layer (Synonyme: Folie, Ebene) können ein Thema weiter in logische Einheiten unterteilen, in denen aus fachlicher Sicht oder aus Sicht eines Verfahrens Geometrieelemente, Objektteile oder Objekte mit gleicher Bedeutung zusammengefasst sind (z.B. Layer innerhalb eines CAD-Themas). Zwischen den Elementen verschiedener Layer können Querbeziehungen bestehen. Layer werden nach pragmatischen Gründen gebildet, um vor allem die rationelle Datenerfassung, die separate Überlagerung und die gemeinsame Darstellung von Teildatenbeständen mit anderen Themen oder Layern für die grafische Präsentation zu unterstützen. Layer bilden eine vertikale Struktur innerhalb eines Themas. Wichtige Kriterien für die Layerbildung sind z.B. die fachliche oder topographische Bedeutung der Inhalte sowie ihre optionale und selektive Nutzung in Karten.

Beispiel FORST-GIS: Layereinteilung des Themas Forstbetrieb (Ebenen)	
Layer	**FBK - Primärgeometrie : Goobjekte Bestand**
1	Grenzpunkte aus der DFK
2	Staatswaldgrenze / Flurstücksgrenze, FoA-Grenzen, digitalisierte Grenzpunkte
3	LKW-befahrbare Waldwege
4	Gewässer
5	Rückewege und Schneisen (Bestandsgrenzen)
6	Rückewege und Schneisen (DISTR-, ABT-, UABT-Grenzen)
7	Sonstige forstwirtschaftliche Grenzen (DISTR, ABT, UABT)
8	Bestandsgrenzen, Bestandsteilflächen (Geoobjekt) → Geolink mit Tabelle Bestand
	FBK - Sekundärgeometrie : Infrastrukturen und Sonderflächen
11	Rückewege, Schneisen, Steige, Einpunktierungen, Nachhiebsreste, Überhälter, u.a..
16	Sonstige Gewässer
22	Schutzwaldflächen
23	Versuchs- und A.R.B.-Flächen; Nassbodenflächen
24	Immissionsflächen
26	FoD-spezifische Sonderflächen
	FBK - Sekundärgeometrie : Signaturen, Texte, Symbole
9	Flächen: Balkensignaturen (BI, NHB, LAT, ...), Nutzungsarten für AW, HG
10	VVJ-Flächen, VVJ-Symbole, Berechtigungsalmen
12	Begleitbänder und Symbole Distriktgrenze, Distriktname
13	Begleitbänder und Symbole Abteilungsgrenze, Abteilungsname
14	Begleitbände Unterabteilungsgrenze, Unterabteilungsname
15	Bestandshochziffern, dringliche Pflegemaßnahmen
17	Texte: Baumarten, Gewässer-, Wege-, Bergnamen, u.a.
20	Zaunlinien, Zaunsymbole
21	Begleitbänder NSG, NR; Symbole: Naturschutzgebiet, Naturwaldreservat
23	Begleitbänder und Symbole WSG; Symbole: Versuchsfläche
24	Begleitbänder FFH/SPA-Gebiete, Symbole FFH/SPA
25	Begleitbänder Sanierungsgebiet; Symbole: Verbauung, Verbauung geplant
30	Begleitbänder FOA-Grenzen
	FBK - Raumbezug und Hintergrund
28	UTM-Gitter
29	Passkreuze und Rahmen der Forstkarten
30	FK5-Rahmen für den Raumbezug, FK5-Blattbezeichnungen
Raster	TK25 Grundriss und Höhenflurkarte

Verschiedene Layer können auch aufgrund von topologischen Beziehungen zwischen Geometrieelementen miteinander verknüpft sein, z.B. wenn die Grenzpunkte in einem Layer von den Grenzen in einem anderen Layer referenziert werden, um die redundante Abbildung der Punkte zu vermeiden.

Beispiel: AutCAD-Zeichnungen = layerstrukturierte Themen.

Geoobjekte - horizontale Struktur von Themen

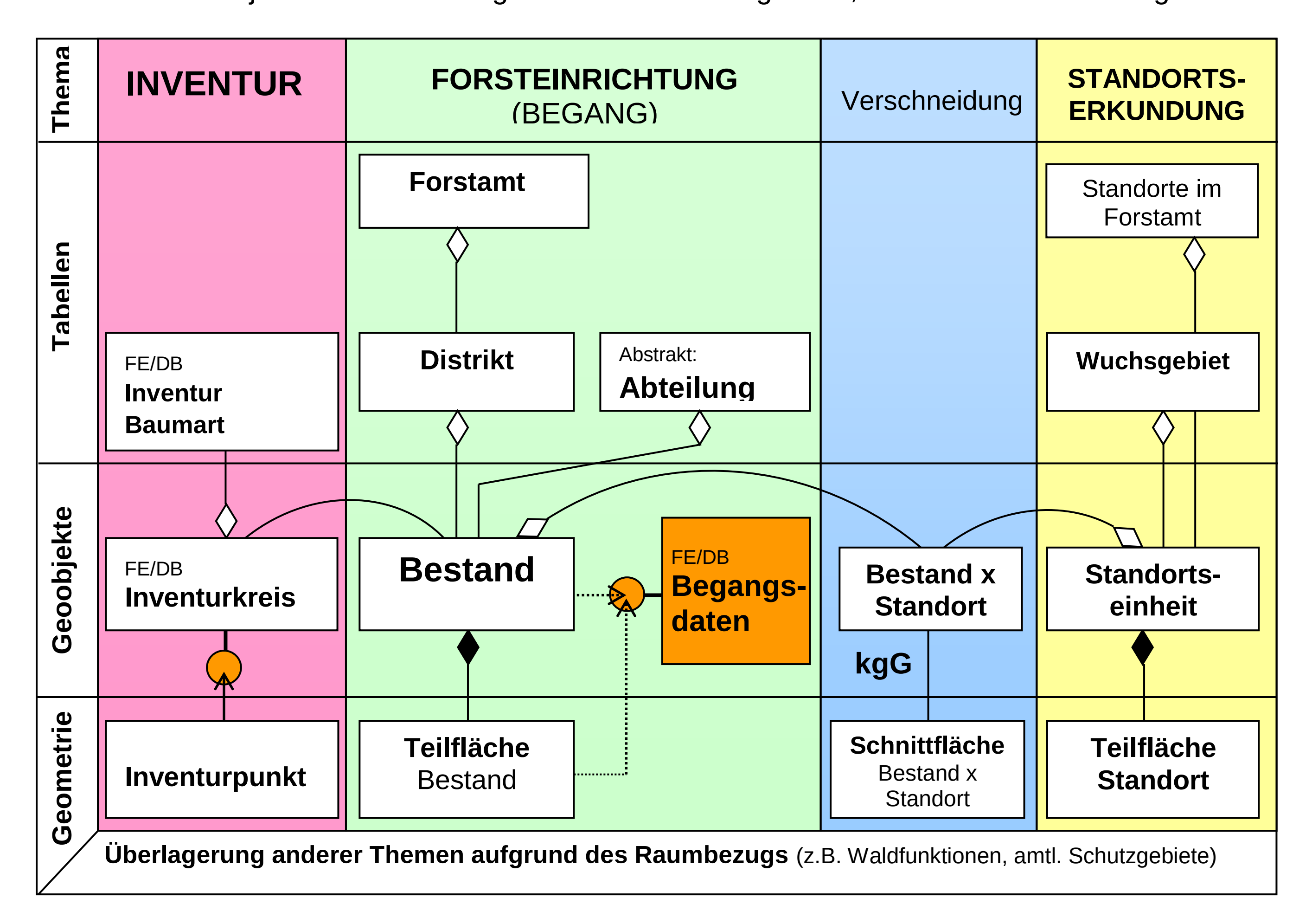
Geoobjekte bilden die horizontale Struktur eines Themas, sie beschreiben raumbezogene Entitäten semantisch und geometrisch. Es sind Aggregationen oder Kompositionen von Objektteilen aus Geometrieelementen und Attributen zu einem „Mosaik".

Geoobjekte sind eindeutig identifizierbare Individuen der modellierten Welt, die als Instanz einer festgelegten (Objekt)-Klasse generiert werden.

Die Geometrieelemente enthalten die vektorielle Information mit dem Geokode und den topologischen Beziehungen, die Tabellen enthalten die attributive Beschreibung und die semantischen Beziehungen eines Geoobjekts.

Wichtige Anwendungen von Objektstrukturen:

- Einzelne, ausgewählte Merkmale in Form von thematischen Karten darstellen
- Räumliche Analysen durch Verschneidungen durchführen
- Objektbezogene Zugriffe ermöglichen, z.B. für die Berechnung von Bestandsflächen
- Selektive objektweise Änderungen der Daten ermöglichen, z.B. für die Fortführung



Das Schema zeigt für die drei Hauptthemen des FORST-GIS: Begang, Inventur und Standorte (unabhängige Kerndatenbereiche), deren Basisgeoobjekte und deren Aggregationen sowie ihre Beziehungen, die für eine Mehrfachnutzung der Geodaten wichtig sind. Die Basisgeoobjekte sind aus fachlicher Sicht die kleinsten logischen Einheiten oder Basisbezugsräume in einem Modell, auf die sich fachliche Aussagen beziehen. Beispiel: ESRI-Shape = einfache Objektstruktur, ohne Topologie.

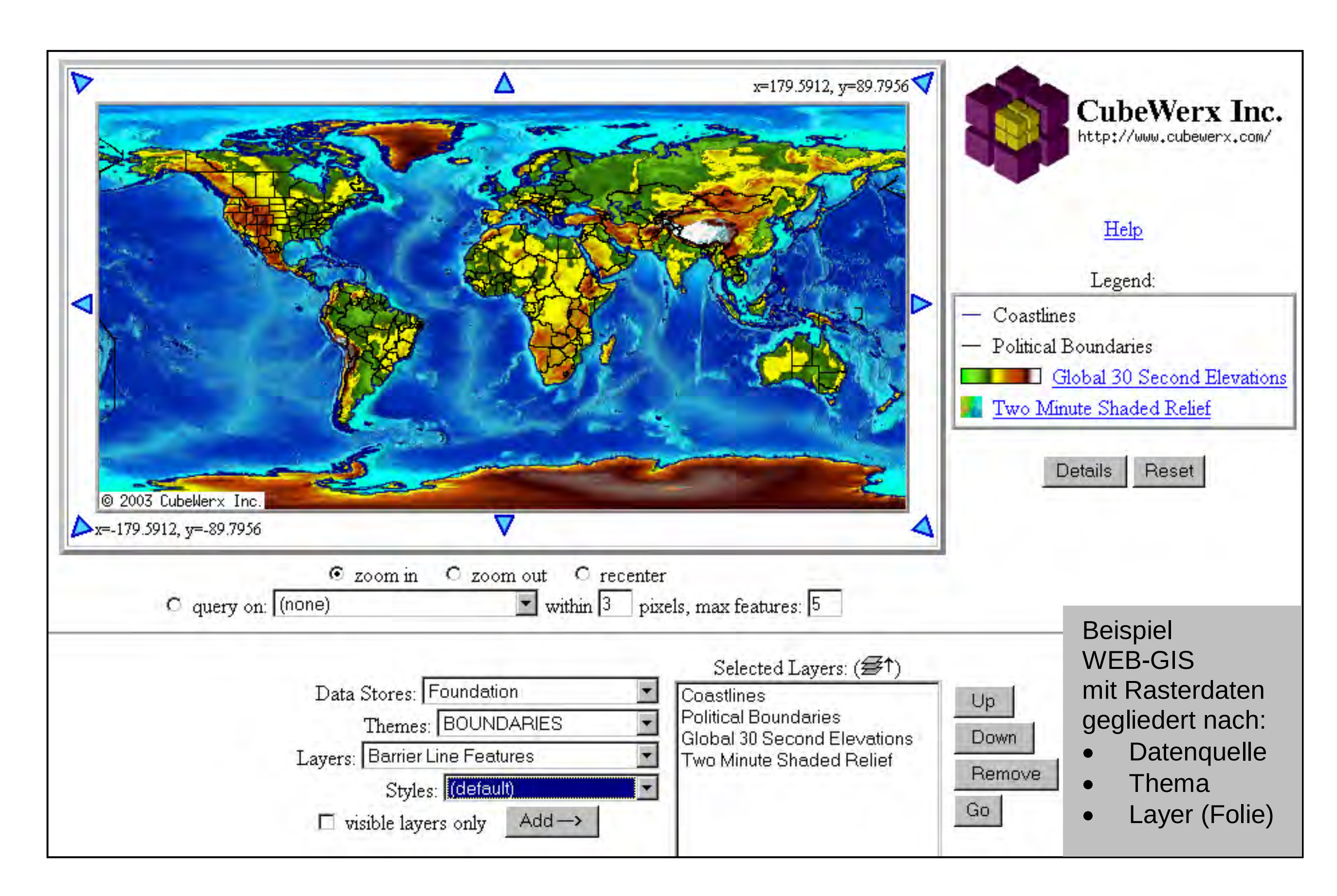
Konzepte für die Organisation von Geodaten

Neben den Möglichkeiten zur Modellierung und Strukturierung von Geoinformation in GIS stehen weitere Konzepte für die Organisation der Geodaten zur Verfügung. So können z.B. Projekte, Datenrahmen oder Views gebildet werden, in denen verschiedene Datenquellen eingefügt (referenziert) werden, um sie zu verwalten sowie ihnen u.a. Parameter für ihre grafische Präsentation oder die Transformation ihres Raumbezugs zuzuordnen. Für die grafische Präsentation (View) können den Geodaten in Projekten, je nach Arbeitsvorgang, unterschiedliche Grafikparameter (Zeichenvorschriften, Styles) zugewiesen werden, z.B. kann für die interaktiv grafische Bearbeitung ein anderer Layout gewählt werden als wir für die Ausgabe als Karte.

Für die Modellierung, zur Verwaltung, und den Zugriff auf Geodaten sind in GIS folgende Konzepte und Datenstrukturen gebräuchlich, die jedoch von einer GIS-Software nicht immer alle unterstützt werden.

- Datenquelle (Data Store, Datenbank)
- **Thema** mit Vektor- oder Rasterdaten, die nur über den Raumbezug mit anderen Themen verbunden sind
- **Layer** für die vertikale Struktur innerhalb eines Themas (Ebene)
- **Geoobjekt** für die horizontale Struktur innerhalb eines Themas (Mosaik, Puzzle)
- **Geometrieelement** als elementare Struktur von Vektordaten
- **Attribute, Tupel** für die Beschreibung thematischer Merkmale
- **Rastermatrix** für Bilddaten mit Pixeln und Grid-Zellen

Beispiel: Datenorganisation in einem WebMapping-Sytem



Beispiel: Datenorganisation in ArcGIS

Projekt: Projekte werden in „Karten“ oder „Szenen“ .mxd|.sxd (Metafiles) organisiert

DataFrame: Ein Projekt muss mindestens ein, kann aber mehrere Dataframes (Datenrahmen, Views) enthalten

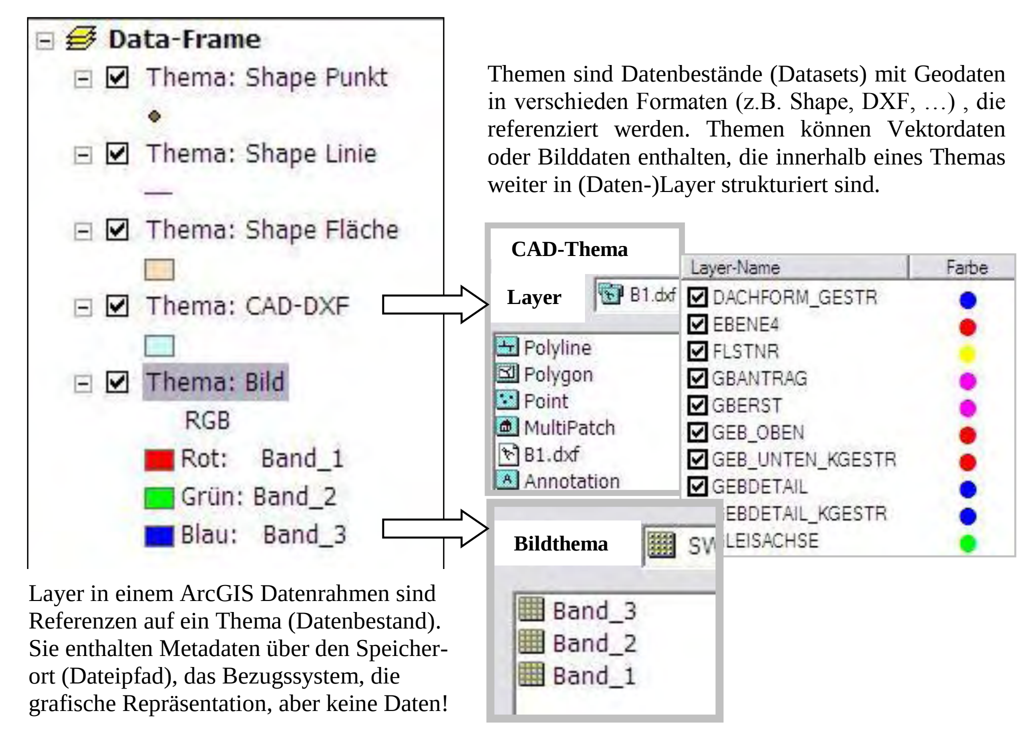
Layer: In einen Dataframe können ein oder mehrere Themen eingefügt (referenziert) werden → diese **Referenz** wird in ArcGIS als „Layer“ bezeichnet. Verwandte Layer können zusätzlich gruppiert werden (GroupLayer)

⇩

Thema = Datenbestand**.** Themen können Vektordaten (Objekte) z.B. Shape, CAD-Daten, Rasterdaten (Bildthemen) oder TIN, GRID enthalten

Layer: Ein Thema kann weiter in **Daten**-Layer unterteilt sein, dies sind logische Einheiten von Daten „Ebenen“ in einem Thema z.B.:
DXF → Layer Featuretype, Layer Inhalt
Bild → Layer Farbkanal (RGB)

<Projekt>.mxd → Inhaltsverzeichnis (table of contents)



Themen sind Datenbestände (Datasets) mit Geodaten in verschieden Formaten (z.B. Shape, DXF, …) , die referenziert werden. Themen können Vektordaten oder Bilddaten enthalten, die innerhalb eines Themas weiter in (Daten-)Layer strukturiert sind.

Layer in einem ArcGIS Datenrahmen sind Referenzen auf ein Thema (Datenbestand). Sie enthalten Metadaten über den Speicherort (Dateipfad), das Bezugssystem, die grafische Repräsentation, aber keine Daten!

Anmerkung:

Die Datenmengen die in einem GIS zur Modellierung und Beschreibung der relevanten Geoinformation über einen raumbezogenen Sachverhalt oder für ein räumliches Phänomen benötigt werden, hängen wesentlich ab von:

- Geometrische Dimension des Modells (2D, 3D)
- Anzahl der Themen und Struktur der Objekte (klein-/großstrukturiert)
- Datentyp (Vektor- oder Rasterdaten)
- Räumliche Ausdehnung (Global | Regional | Lokal)
- Räumliche Auflösung*) (Übersicht | Rahmenplan | Detailplan)

**) Mit Auflösung wird die kleinste beschreibbare Einheit bezeichnet, die geometrisch und thematisch modelliert wird.*

4 Semantisches Datenmodell (konzeptioneller Entwurf)

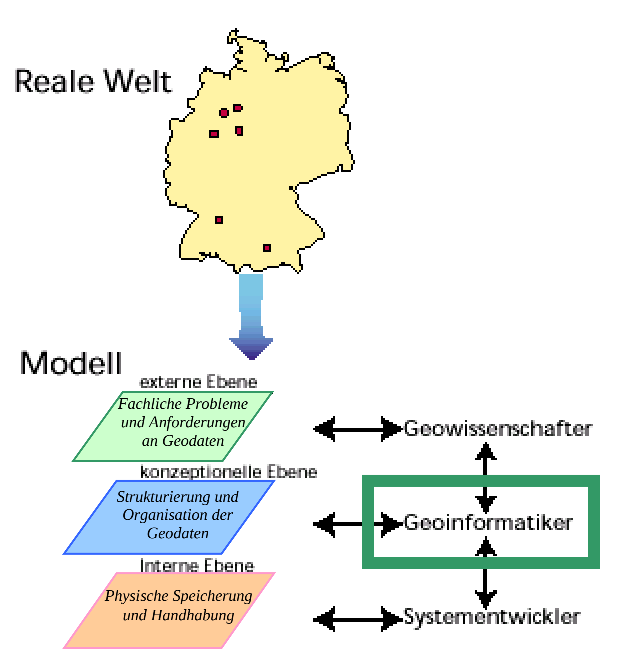
Das Vorgehen beim konzeptionellen Entwurf von Geodatenmodellen wird am Geoobjekt aufzeigt, das die komplexeste Modellierungskomponente für Geodaten darstellt, die dann benötigt wird, wenn Geoinformation objektbezogen verarbeitet werden soll. Geoobjekte sind aus konzeptioneller Sicht nach fachlichen Kriterien gebildete Abstraktionen der Entitäten des abzubildenden Themenbereichs. Sie erweitern die Einsatzmöglichkeiten der Datenbank besonders im Hinblick auf thematische und analytische Auswertungen, die eine hohe Logik im Datenmodell erfordern. Für ihre Implementierung müssen sie mit den verfügbaren Datenstrukturen einer GIS-Software in ein „realisierbares“ logisches Schema überführt werden. Aus technischer Sicht kann es dabei erforderlich sein, fachliche Elementarobjekte im logischen Modell weiter in Subobjekte zu zerlegen oder auf nicht abbildbare Merkmale zu verzichten.

Ein **Geobjekt** ist ein auf einen räumlichen Ausschnitt der Erde bezogenes reales oder gedankliches Objekt, das hinsichtlich seiner

- räumlichen Lage und Form **Geometrie**
- Lagebeziehungen zu anderen Geoobjekten **Topologie**
- fachlich relevanten Eigenschaften **Thematik**
- zeitlichen Veränderung Dynamik

gegenüber anderen Geoobjekten unterschieden wird (Streit).

Fachanforderung (reale Welt)
Anforderungsanalyse
Konzeptioneller Entwurf (UML, ERM) Semantisches Datenmodell
Auswahl eines Herstellersystems
Logischer Entwurf (Datenstruktur, DDL) Logisches Datenmodell
Physischer Entwurf (Implementierung)



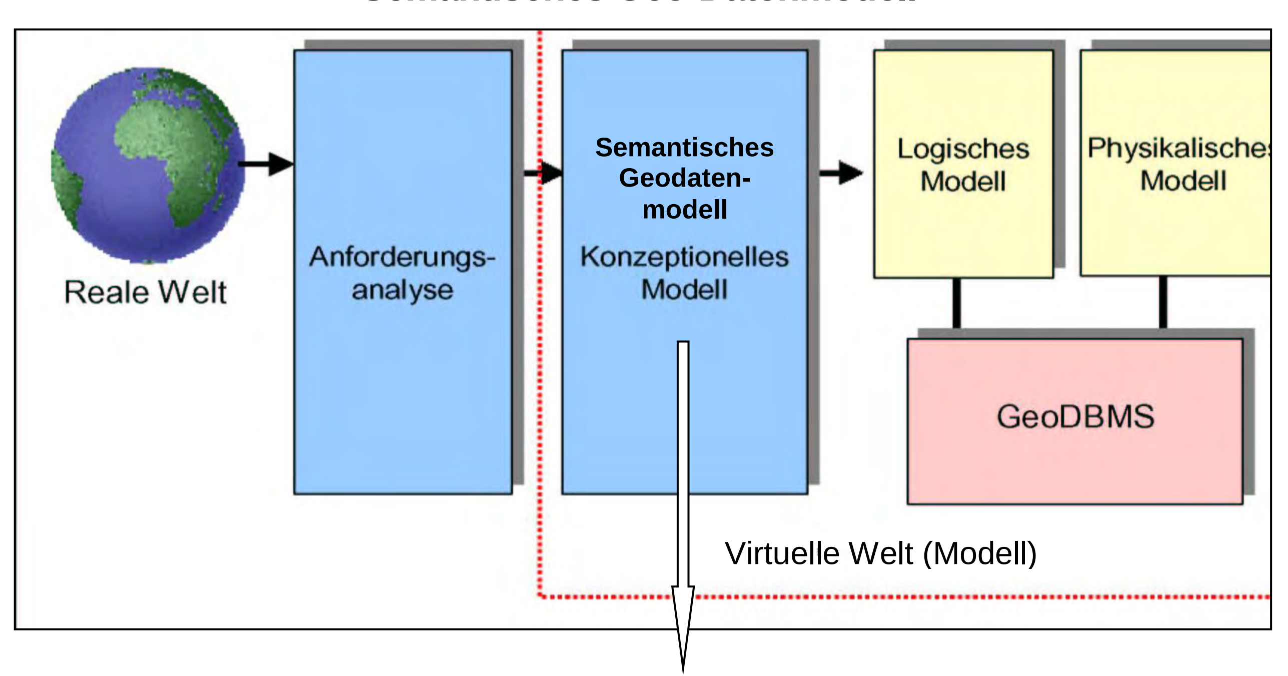
Definition Geoobjekt nach Bill/Fritsch

Ein Objekt im Sinne eines GIS (**Geoobjekt**) ist eine **konkrete** physisch, geometrisch oder begrifflich begrenzte **Einheit** der Natur und besitzt eine **individuelle Identität**

Jedes (Geo)Objekt repräsentiert ein **Unikat in der realen Welt**, das jedoch einer bestimmte Objektklasse zugeordnet werden kann.

(Hinweis: Unikate → Unterschied zwischen Modell der realen Welt zu einem Modell einer virtuellen Spielwelt)

Semantisches Geo-Datenmodell



Semantisches Geodatenmodell nach Meng

Reale Welt	primäres Modell **Geodatenmodell**		+ Styles	sekundäres Modell **Visuelles Modell**	tertiäres Modell
	Objektebene (**Feature** Level)		**Modell**ebene	**Darstellung**sebene (**Map** Level)	**Denk-**ebene
Reale Welt Objekte (Entität)	räumliche Daten	Geoobjekte	Ontologie	(Mehrfach-) Repräsentation	Benutzer betrachtet die Karte (Präsen-tation)
	Geometrie (CRS)	Geometrie + Sachdaten	**Geoobjekte** mit ihren **Beziehungen**	**Graphik** (in Abhängigkeit der Styledescription für Features)	
Haus Straße Baum	Fläche Linie Punkt	+ Hausnummer Straßenname Baumkataster-nummer	Häuser an einer Allee	kartographisch ausgestaltete Darstellung (BS oder Papier) skizzenhafte Bildschirmgrafik (Topogramm) 3D-Darstellung	z.B. Stadt-planer
Geoobjekte und ihre Beziehungen im primären Modell **repräsentieren** die Entitäten (Objekte) der realen Welt				Graphikelemente im sekundären Modell **visualisieren** die Geoobjekte des primären Modells	

Die Verarbeitung von Geoinformation beruht auf vier Komponenten: Geometrie, Sachdaten, Topologie und Grafik. Die **Geometrie** bildet die räumliche Lage und Ausdehnung von Objekten der realen Welt (Entitäten), wie z. B. Bäume, Straßen und Häuser, näherungsweise als Punkte, Linien oder Flächen ab. Fügt man der Geometrie zusätzlich alphanumerische Informationen in Form von **Attributen** hinzu, wie z. B. Hausnummer, Straßenname oder Baumkatasternummer, entstehen **Geoobjekte.**

Bringt man die einzelnen Geoobjekte und ihre Beziehungen (attributive und **topologische)** in einen Gesamtzusammenhang (Ontologie), so entsteht das **Geodatenmodell** (primäre Modell) als Abbild der realen Welt.

Grundgedanken der objektstrukturierten Modellierung

Wegen der Allgemeinheit des Wortes „Objekt“ wird auch der Begriff „objektorientiert“ viel benutzt und ist bei unvorsichtiger Verwendung stets in Gefahr ein Allgemeinplatz zu werden. Vorab sollen deshalb die Begriffe Objekt und objektorientiert kurz in ihrer Bedeutung umrissen werden.

Unter einem „**Objekt**“ wird hier ein allgemeines technisches Konzept verstanden, mit dem man die realen Objekte der Umwelt, die zur Unterscheidung als Entitäten bezeichnet werden, so abstrahieren kann, dass sie von einem Rechner als logische Informationseinheiten verarbeitet und verwaltet werden können.

→ Objekt ist ein allgemeines technisches Konzept

Der Begriff „**objektorientiert**“ ist von der Informatik fest belegt und bezeichnet ganz spezielle Konzepte für die Erstellung von Programmsystemen oder für die Modellierung von Datenbanken. Die heute übliche Verwendung von Geoobjekten entspricht nur in Teilen diesen Prinzipien der Objektorientierung, die meist nur auf der konzeptionellen Ebene vollständig eingesetzt werden. Die nachfolgend dargestellte Verwendung von Geoobjekten wird deshalb zur Unterscheidung von den streng objektorientierten Konzepten der Informatik als „**objektstrukturiert**“ bezeichnet.

→ Ein Geoobjekt ist eine logische Informationseinheit für die objektstrukturierte Modellierung räumlicher Entitäten

Als strategische Referenz für den objektstrukturierten Entwurf von Geoobjekten werden zunächst die grundlegenden Konzepte der Objektorientierung in Anlehnung an die Veröffentlichungen der *DOOD*.

DOOD = International Conference on Deductive and Object-Oriented Database Systems

kurz zusammengestellt, um daraus das Vorgehen für die Konzeption und den Entwurf von Geoobjekten abzuleiten.

Die Objektorientierung basiert auf dem Grundgedanken:

„Objektorientierte Systeme betrachten und modellieren die Entitäten der realen Welt als eine Kollektion von kooperierenden und untereinander in Beziehung stehenden wohlunterscheidbaren Einheiten, den Objekten. “

Dieser Ansatz hat viele Gemeinsamkeiten mit dem Wesen gegenständlicher Systeme, wobei die Objekte den Systemelementen und die Relationen den Wirkungsbeziehungen entsprechen.

Der Objektorientierung liegen vier wesentliche Konzepte zugrunde, die:

- **Autonomie** (Kapselung von Eigenschaften und Methoden im Objekt)
- **Abstraktion** und **Klassifikation** (Zusammenfassung gleichartiger Objekte)
- **Vererbung** (Taxonomie, Generalisierung, Spezialisierung)
- **Polymorphismus** (overriding, overloading von Methoden)

Objektorientierung wird eingesetzt sowohl als

- analytische Technik für die objektorientierte **Systemanalyse** und
- analytische Technik für den objektorientierten **Systementwurf** sowie
- konstruktive Technik für die objektorientierte **Systemkonstruktion**

für Programme und Datenbanken.

Unterschied Datenbank / Programm

Datenbank- und Programmsysteme verfolgen grundsätzlich andere Zielsetzungen. Während ein Programmsystem aus einer exakt festgelegten Menge von Objekten für eine bestimmte Aufgabe konstruiert wird, muss eine Datenbank so beschaffen sein, dass sie einen großen Objektvorrat mit nicht absehbarer Extension persistent vorhalten kann, aus dem Objekte nach Bedarf ausgewählt werden können, und der für die gemeinsame Nutzung durch viele Applikationen zur Verfügung steht. Für Datenbanken steht insbesondere das Konzept der Kapselung in einem prinzipiellen Konflikt zu ihren Erfordernissen; für Datenbankobjekte ist deshalb der Kapselungsbegriff „liberal" zu behandeln (*Dittrich 1997*).

Konzepte der OO-Programmierung	Konzepte der OO-Datenmodelle
• **Programmobjekte** • Klassen mit gekapselten Eigenschaften und Methoden • Vererbung (Generalisierung, Spezialisierung) • Polymorphismus und spätes Binden • Metaklassen (Interface)	• **Datenobjekte** • Klassen mit Attributen • Objektidentität und Persistenz • Vererbung (Taxonomie) • Beziehungen zwischen Objekten (Assoziation, Aggregation) • Integritätsbedingungen

Damit ein Datenbanksystem als objektorientiert (OODBMS) bezeichnet werden kann, sollte es nach Vorgabe der *DOOD* mindestens die nachfolgenden

13 Grundeigenschaften (Detailkonzepte)

erfüllen, die man quasi als eine Checkliste für objektorientierte Datenbanksysteme ansehen kann.

1. Klassenkonzept und Definierbarkeit von Klassen (Classes)
2. Objektidentität, wird durch System gewährleistet (Object Identity)
3. Zusammengesetzte Objekte (Complex Objects)
4. Klassenhierarchie und Vererbung (Class Hierarchies and Inheritance)
5. Erweiterbarkeit um neue Datenstrukturen und Typen (Extensibility)
6. Dauerhafte Verwaltung von Datenelementen (Persistence)
7. Paralleler Zugriff, Mehrfachbenutzbarkeit, Nebenläufigkeit (Concurrency)
8. Optimierung der internen Datenorganisation (Secondary Storage Management)
9. Wiederherstellung im Störfall (Recovery)
10. Einkapselung für Zustand und Operationen (Encapsulation)
11. Überschreiben, Überladen und spätes Binden (Overriding, Overloading and Late Binding)
12. Sprache zur Formulierung von Operationen, Berechnungsvollständigkeit (Computational Completness)
13. Abfragesprache für interaktive Zugriffe (Ad Hoc Query Facility)

Dittrich ergänzt:

„Ein OODBMS muss ein **objektorientiertes Datenmodell** aufweisen, was im Einzelnen die Unterstützung der oben genannten Detailkonzepte bedeutet, und

es muss außerdem alle funktionalen Eigenschaften eines klassischen Datenbanksystems haben".

Objekte

Objekte sind Abstraktionen der statischen Struktur von Entitäten mit deskriptiven Eigenschaften und Assoziationen. Sie sind der Kerngedanke objektorientierter Systeme. Objekte können als allgemeines technisches Konzept der Informatik aufgefasst werden, das folgende wesentliche Inhalte hat:

- Ein Objekt ist eine Abstraktion einer Entität mit definierten Grenzen und einer **definierten Identität**; jedes Objekt ist ein Unikat.
- Klassen werden eingeführt, damit die Vielzahl der Objekte beherrschbar bleibt. Klassen sind Baupläne (Schablonen) gleichartiger Objekte, die gemeinsame Eigenschaften und Verhaltensmerkmale vorgeben und bezeichnen.
- Objekte existieren unabhängig von den Werten ihrer Eigenschaften; während ihr Zustand, die Werte, sich ändern können, bleibt ihre Identität unveränderlich erhalten.
- Bei Objekten ist streng zwischen Gleichheit und Identität zu unterscheiden, zwei Objekte sind dann gleich (Klone), nicht identisch, wenn sie zu einem Zeitpunkt in allen ihren Eigenschaften übereinstimmen
- Objekte werden ausschließlich als Instanzen ihrer Klassen erzeugt.
- Objekte können zu komplexen Objekten aggregiert werden.
- Neue Klassen können als Generalisierung oder Spezialisierung (Spezialisierung vom Allgemeinen zum Besonderen) einer Klasse gebildet werden (Vererbung).

Für die Objektorientierung im Sinne der Informatik kommen zusätzliche Detailkonzepte hinzu, die das allgemeine technische Konzept Objekt weiter spezifizieren, wie z.B. die Implementierung von problemorientierten Operationen als Methoden in den Objekten, entsprechend den geforderten Verhaltensmerkmalen.

Für Sachverhalte mit komplexen Entitäten, wie sie in Geodaten vorkommen, ist die Nutzung des Objektkonzepts für die Strukturierung der Daten auch ohne die vollständige Realisierung aller Detailkonzepte der Objektorientierung sinnvoll. Es muss dabei eine Adaption mit den Mitteln der verfügbaren, meist nicht vollständig objektorientierten GIS- und Datenbankprodukten erfolgen. Solche Datenmodelle, die sich nur das technische Konzept Objekt - mit Klassen, Objekten, komplexen Objekten und ihrer Taxonomie - stützen, bezeichnet man als strukturell objektorientiert oder kurz als **objektstrukturiert**. Eine Datenbank enthält dann eine Menge von persistenten Datenobjekten, die von Klassen abstammen, und ihre Beziehungen.

Einschub:

Metazeichen nach Backus Naur	
Regel ::= <Verknüpfung von Ausdrücken>	
::=	Definition, trennt linke von rechter Seite einer Regel
|	Alternative, exklusive or-Verknüpfung (entweder/oder)
,	Aufzählung, and-Verküpfung
...	Bereich (von, bis)
[ ]	optional 0 **|** 1
{ }	Iteration, 1...n (beliebig oft, jedoch mindestens 1 mal)
( )	Logische Gruppierung von Teilausdrücken, die zusammengehören

Geoobjekte

Geoobjekte sind Abstraktionen räumlicher Entitäten → Erweiterung des Objektbegriffs

DOOD: An **object** in an OODB can be regarded as a quadruple:

<OID,class,value,state>

and a **class** as an 11-tuple:

< ***class-name****, member value type, member state type, {member messages}, {member methods}, class value, class state, {class messages}, class value type, class state type, {class methods}* >

Sichtbare Objekteigenschaften:

< OID, Klasse, {Attribute} >

Sichtbare Klasseneigenschaften:

< Klassenname, {Klassenattribute/-typen}, {Klassenoperationen} >

Geoobjekte werden für die Abstraktion raumbezogener Entitäten eingesetzt. Für ihre Modellierung sind geometrische Datenstrukturen, die hier als Geometrieelemente bezeichnet werden erforderlich, zusätzlich zu den üblichen Attributen zur Beschreibung thematischer oder temporaler Sachverhalte.

Damit ergeben sich folgende sichtbaren Eigenschaften von (elementaren) Geoobjekten

Geoobjekt ::= < OID, Klasse, {Geometrieelemente}, {Attribute} >

Die sichtbaren Eigenschaften der Geoobjekte kann man in Anlehnung an ihren primär metrischen, form- und lagebeschreibenden Charakter als Konfiguration der Geoobjekte bezeichnen. Aus dem erweiterten Aufbau folgt, dass für Geoobjekte eine geometrische und eine thematische Modellierung erforderlich sind. Ein Geoobjekt besteht mindestens aus einem Geometrieelement und einem Attribut, d.h. ein Attribut oder Geometrie-element allein, kann kein Geoobjekt sein.

Objektidentität (OID)

Ein wichtiger Aspekt von Objekten ist die Wahrung ihrer **Identität**, die durch einen geeigneten Object Identifier (OID) zu realisieren ist, für den gilt: „An object identifier (OID) **is a system-supported concept, seperate from the value of an object**, to support object identity within the database. To meet the requirements of object identity, an OID has to be **system-wide unique**, has to remain **unchanged over the lifetime** of the object, and **cannot be reused** for other objects“. Aus dieser Definition resultieren folgende Anforderungen, die ein OID erfüllen muss:

- ein OID ist **systemweit eindeutig**
- ein OID ist **unveränderlich** während der gesamten Lebensdauer des Objekts
- ein OID wird nach dem Untergang des Objekts **nicht wiederverwendet**

Sofern jedes Objekt eine solche OID enthält, unabhängig von seiner Komplexität, ist für die vollkommene Unabhängigkeit seines Zustands (Werte) und seiner individuellen Existenz gesorgt.

Die OID ergibt die systemweite („anonyme“) Identität eines Geoobjekts, sie entspricht einer „Personalausweisnummer“. Für die logische Ansprache von Geoobjekten sind neben der OID oft weitere Attribute erforderlich, die eine eindeutige Bezeichnung beinhalten, sie entsprechen den „Personalien“ (fachliche Identität). Der Raumbezug ist die Adresse (Geokode), der angibt, wo sich das Geoobjekt befindet.

Unified Modeling Language (UML) - Auszug

Übersicht über die Elemente des Klassendiagramms

UML = Unified Modeling Language

UML = Objektorientierte Sprache und Notation zur
- **Spezifikation**
- **Konstruktion**
- **Visualisierung**

von Modellen für Softwaresysteme

UML = keine Entwurfsmethode

UML = sehr komplex und enthält eine Vielzahl einzelner Konstrukte und Elemente. Für die Beschreibung von Informationssystemen reicht das Klassendiagramm aus.

UML = OMG-Standard (Object Management Group), **ISO-Norm 19501**

Klassendiagramme enthalten die statische Struktur eines Datenmodells, die sich - unter Anwendung der Regeln für die Anforderungsanalyse - aus den Fachvorgaben ergibt. Klassendiagramme zeigen eine Menge statischer Modellelemente, vor allem die Klassen und ihre Beziehungen (Assoziationen).

Basiselemente für ein Klassendiagramm
- Klasse (Objektklasse)
- Attribute, Operationen
- abstrakte Klasse
- Interface (Schnittstelle)
- Objekt (Instanz), Instanziierung

Beziehungselemente im Klassendiagramm
- Assoziation, Multiplizität
- Aggregation, Komposition (Hierarchie)
- Vererbung (Generalisierung, Spezialisierung)

UML versus ERM

UML ist ein plattformunabhängige, objektorientierte Notation für das konzeptionelle Schema, die nicht das logische Schema vorwegnimmt, sie ist offen für alle Datenbankmodelle

ERM ist stark auf das relationale Modell ausgerichtet, es nimmt damit das logische Schema weitgehend vorweg.

Die Unified Modelling Language hat sich in der GIS-Welt in den letzten Jahren für die Darstellung konzeptioneller Schemata etabliert.

(OGC, Document 99-010: „to add UML as a mandated Standard for use in OGC specifications“)

Literatur:

Oestereich B. (2001)	*Die UML-Kurzreferenz für die Praxis, Oldenbourg Verlag München Wien*	*http://www.oose.de/uml*
Balzert H. (2005)	*UML 2, kompakt mit Checklisten Spektrum, Akademischer Verlag*	
Rumbaugh J., Booch G., Jacobson I. (1999)	*Das UML-Benutzerhandbuch, Addison-Wesley Verlag*	

Basiselemente für das Klassendiagramm

Klasse (Objektklasse) Notation und Beispiel	**Eine Klasse ist die Definition der Attribute, Operationen und der Semantik einer Menge von Objekten. Alle Objekte einer Klasse (Extension) entsprechen dieser Definition. Eine Klasse enthält die Beschreibung der Struktur und des Verhaltens von Objekten, die sie erzeugt oder die mit ihr erzeugt werden können.**
Klassen werden durch Rechtecke dargestellt, die entweder nur den Namen der Klasse tragen oder zusätzlich auch Attribute und Operationen. Die Rubriken – Klassenname, Attribute, Operationen – werden durch horizontale Linien getrennt. **Klasse** **Klasse** / attribute / operationen Klassennamen beginnen mit einem Großbuchstaben sind Substantive im Singular (fettgedruckt) Bei Bedarf können Attribute und Operationen weiter spezifiziert werden. **Klasse** / attribut: Typ=Initialwert {Zusicherung} / operation (Argumente) {Zusicherung}	**Haus** **Student** / name, vorname / immatrikulieren (), exmatrikulieren () **Kreis** / radius {radius>0}, mittelpunkt: Point=(10,10) / setPosition (pos: Point), setRadius (neuerRadius), anzeigen ()

Attribut Notation und Beispiel	**Ein Attribut ist ein (Daten-)Element, das in jedem Objekt einer Klasse gleichermaßen enthalten ist und von jedem Objekt mit einem individuellen Wert repräsentiert wird. Attribute haben keine Identität, sie sind Teile von Objekten.**
Attribute werden mindestens mit ihrem Namen aufgeführt und können zusätzliche Angaben zu ihrem Datentyp, einem Initialwert, Eigenschaftswerte und Zusicherungen enthalten. Attribute beginnen mit Kleinbuchstaben, Eigenschaftswerte und Zusicherungen stehen in geschweiften Klammern.	Sichtbarkeit attribut: Typ=Initialwert {Zusicherung} + radius: Integer=25 {radius>0} ↑ Sichtbarkeit (+ public, # protected, - private)

Operation Notation und Beispiel	**Operationen sind Dienstleistungen, die von einem Objekt angefordert werden können, sie werden beschrieben durch ihre Signatur (Operationsname, Parameter) : Typ. Eine Methode implementiert eine Operation.**
Eine Operation hat innerhalb einer Klassendefinition eine eindeutige Signatur, die sich aus dem Namen der Operation, den formalen Parametern und einem ev. vorhandenen Rückgabewert (Funktionsergebnis) zusammensetzt. Die Namen von Operationen und Parametern beginnen mit einem Kleinbuchstaben. Operationsnamen für den Zugriff auf Attributwerte beginnen oft mit set oder get . Die konkreten Übergabewerte von Parametern bezeichnet man als Argumente.	Sichtbarkeit operation (parameter) : Rückgabetyp + setRadius (neuerRadius) : Double Parameter kann man bei Bedarf weiter spezifizieren (Art parameter : Parametertyp=Standardwert) ↑ Art (in, out, inout) + setPosition (in x : Integer = 1, in y : Integer =1) : void + berechneFlaeche (out fla : float) : Double

Abstrakte Klasse

Eine abstrakte Klasse ist immer eine Oberklasse, von der niemals Instanzen (Objekte) erzeugt werden.

Eine abstrakte Klasse ist unvollständig und bildet die Basis für Unterklassen.

Eine abstrakte Klasse die keine Unterklassen hat, ist überflüssig.

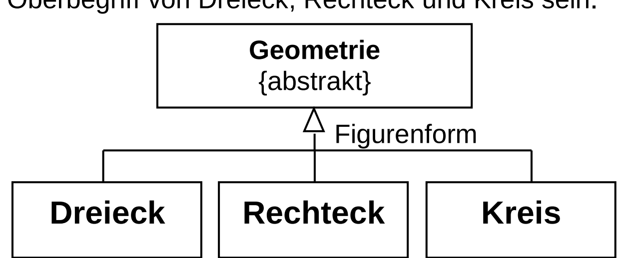
Notation und Beispiel

Eine abstrakte Klasse wird wie eine normale Klasse dargestellt, unter den Klassennamen steht jedoch der Stereotyp {abstrakt}

Klasse {abstakt}

Klasse {abstakt}
attribute
operationen

Abstrakte Klassen repräsentieren häufig einen Allgemeinbegriff. So kann Geometrie ein abstrakter Oberbegriff von Dreieck, Rechteck und Kreis sein.



Eine **Oberklasse** (Basisklasse) ist die **Verallgemeinerung** ausgewählter Eigenschaften (Generalisierung) ihrer Unterklassen

Eine **Unterklasse** (Subklasse) ist die **Spezialisierung** einer Oberklasse, sie erbt alle Eigenschaften ihrer Oberklasse

Interface

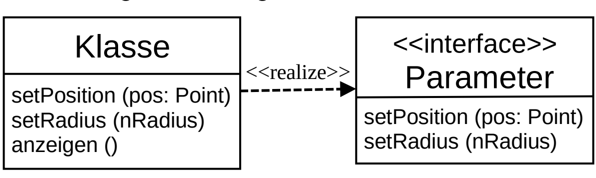
Ein Interface (Schnittstelle) ist eine abstrakte Klasse, die ausschließlich abstrakte Operationen (Methoden) definiert. Es sind Spezifikationen des externen Verhaltens von Klassen und enthalten eine Menge von Signaturen für Operationen, die implementiert werden müssen, falls die Schnittstelle in einer Klasse genutzt wird.

Notation und Beispiel

Eine Schnittstellenklasse wird wie eine normale Klasse dargestellt, über den Klassennamen steht jedoch der Stereotyp <<interface>>

Ein Klasse, die eine Schnittstelle implementiert, muss alle spezifizierten Operationen realisieren. Eine Klasse kann mehrere Schnittstellen implementieren und darüberhinaus weitere Eigenschaften und Mehtoden enthalten.

Zwischen Klasse und einer Schnittstelle besteht eine Realisierungsbeziehung.



Objekt (Instanz)

Ein Objekt ist eine im System konkret vorhandene Einheit.
Jedes Objekt ist ein Exemplar seiner Klasse.
Ein Objekt erhält durch Attributwerte repräsentierte individuelle Merkmale, deren Struktur in der Klasse definiert ist.
Ein Objekt kann die in der Klasse definierten Operationen nutzen

Notation und Beispiel

Objekte werden durch Rechtecke dargestellt, die entweder nur ihren Namen tragen, aber zusätzlich auch den Namen ihrer Klasse oder auch Werte bestimmter oder aller Attribute enthalten können. Werden Attributwerte angegeben, wird das Rechteck durch eine horzonztale Linie in zwei Rubriken geteilt.

Objektnamen = Exemplarnamen, sie werden unterstrichen und beginnen mit Kleinbuchstaben.

objekt

objekt: Klasse
attributname = wert

neuerKreis: Kreis
radius = 25 Mittelpunkt = (10,10)

Instanziierung

Klassen-Objekt-Beziehungen sind Abhängigkeitsbeziehungen.

Objekte sind immer Instanzen einer Klasse.

Notation und Beispiel

Instanziierungsbeziehungen werden durch einen gestrichelten Pfeil dargestellt, wobei das Objekt auf seine Klasse zeigt.

Kreis ◄- - - - <<instance of >> - - - - neuerKreis

Beziehungselemente für das Klassendiagramm

Im Klassendiagramm können drei wichtige Beziehungstypen von Objekten dargestellt werden, die eine zunehmend engere „Verwandtschaft“ der Objekte wiedergeben:

- Assoziation allgemeine Beziehung, kann zwischen allen Klassen bestehen
- Aggregation Hierarchie, Ganzes-Teile Beziehung
- Vererbung zwischen verwandten Klassen, die spezielle erbt von der allgemeinen

Assoziation, Multiplizität, Kardinalität

Eine Assoziation beschreibt eine Beziehung (Verknüpfung) zwischen den Objekten verschiedener Klassen, aber auch zwischen Objekten derselben Klasse.

Die Multiplizität einer Assoziation gibt an, mit wie vielen Objekten der assoziierten Klasse ein Objekt in Beziehung stehen kann.

Die Kardinalität ist die konkrete Anzahl der in Beziehung stehenden Objekte. Die Multiplizität ist der Bereich erlaubter Kardinalitäten

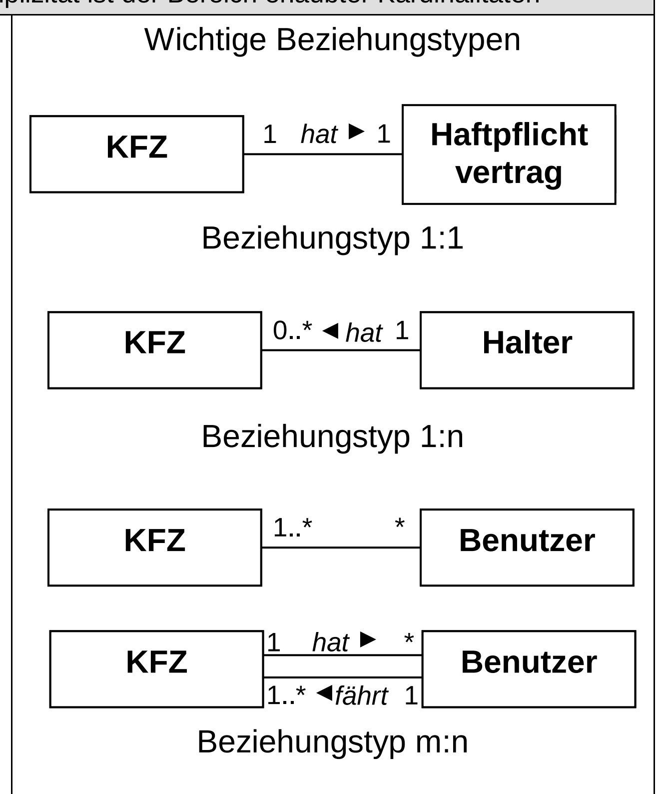
Notation und Beispiel

Eine Assoziation wird durch eine Linie zwischen den beteiligten Klassen dargestellt. An den jeweiligen Enden kann die Multiplizität der Beziehung angeschrieben werden.

Die Multiplizität gibt den Bereich oder Beziehungstyp von Klassen an. Beispiele für **Multiplizitätsangaben**:

1		genau eins
0..1		null oder eins
0..4		zwischen null und vier
3,7		genau drei oder sieben
1..*		größer oder gleich eins
0..*	*****	größer oder gleich Null (**Standard**, wenn Angabe fehlt)

Beziehungen können mit einen Namen (kursiv) versehen werden, der beschreibt worin oder warum diese Beziehung besteht.
Damit die Klassen und die Beziehung in richtiger Reihenfolge gelesen werden, kann neben den Beziehungsnamen ein kleines ausgefülltes Dreieck in Leserichtung gezeichnet werden.
An den Enden der Linie können zusätzlich Rollennamen angegeben werden, Rollennamen beginnen mit einem Kleinbuchstaben.



Vererbung, Generalisierung, Spezialisierung

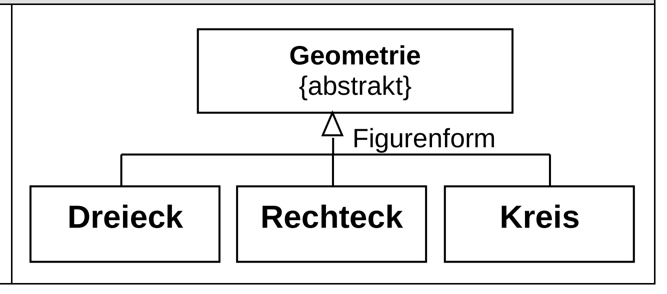
Vererbung ist ein Konzept objektorientierter Systeme für Beziehungen zwischen Ober- und Unterklassen, wodurch Attribute und Operationen der Oberklasse auch den Unterklassen zugänglich gemacht (vererbt) werden.

Eine Generalisierung bzw. Spezialisierung ist eine taxonomische Beziehung zwischen einer allgemeinen und speziellen Klasse bzw. umgekehrt, wobei die speziellere weitere Eigenschaften hinzufügt und sich kompatibel zur allgemeinen verhält.

Notation und Beispiel

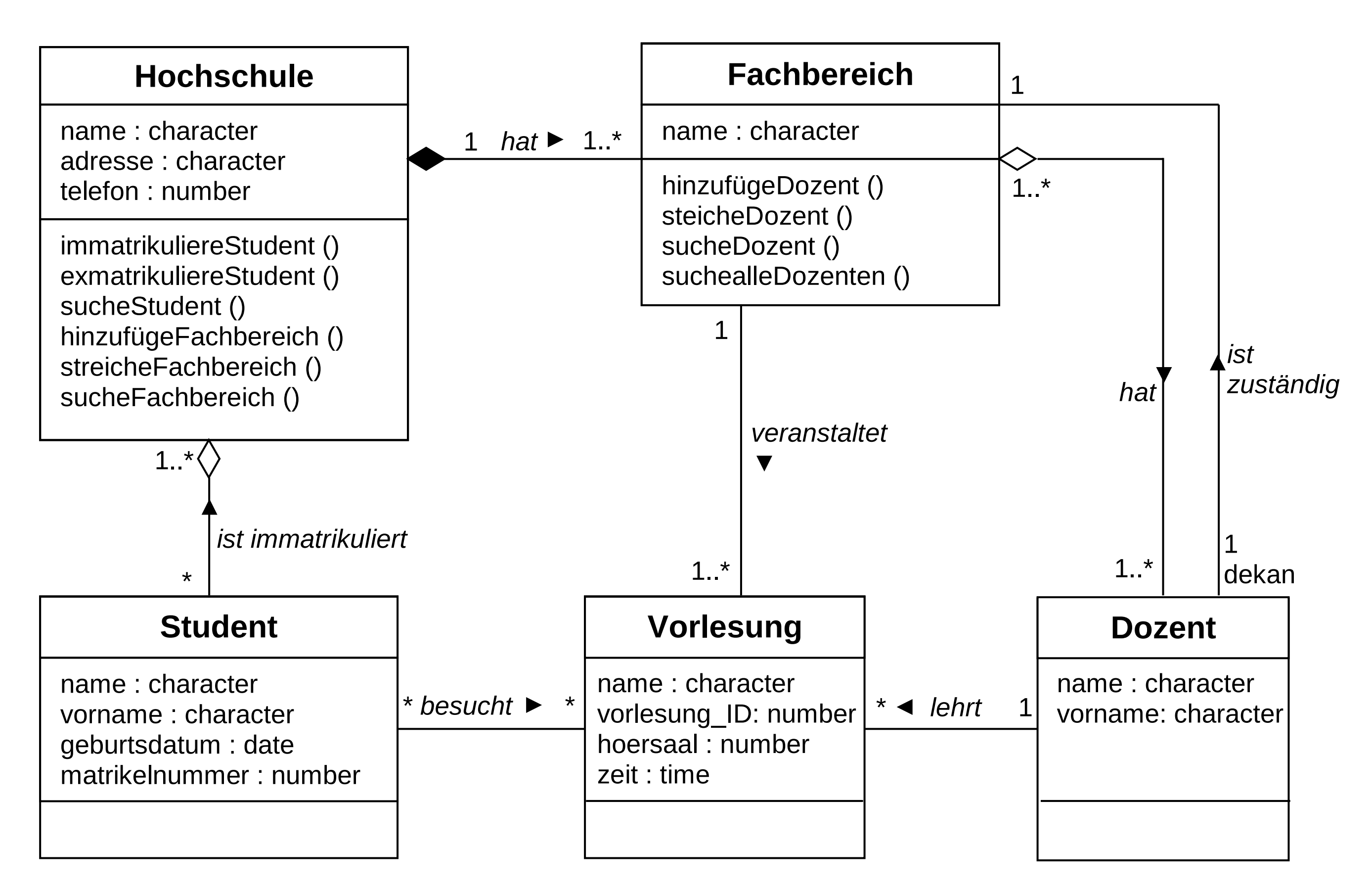
Die Vererbungsbeziehung wird mit einem nicht ausgefüllten Pfeil dargestellt, wobei der Pfeil von der Unterklasse zur Oberklasse (vom Speziellen zum Allgemeinen) gezeichnet wird.

Erfolgt die Spezialisierung einer Oberklasse aufgrund eines einheitlichen Charakteristikums, das als Diskriminator bezeichnet wird, so wird dieses an den Pfeil geschrieben.



Aggregation Komposition Notation und Beispiel	Eine Aggregation ist eine Beziehung, bei der die beteiligten Klassen eine Ganzes-Teile-Hierarchie darstellen, d.h. eine Aggregation gibt an, wie sich ein Ganzes aus seinen Teilen logisch zusammensetzt. Aggregation heißt: „ .. besteht logisch aus ...“ Eine Komposition ist eine strenge Form der Aggregation, bei der die Teile vom Ganzen existenzabhängig sind, d.h. wird das Ganze gelöscht, werden auch alle Einzelteile gelöscht. Wird ein Einzelteil gelöscht, bleibt jedoch das Ganze erhalten, sofern es noch weitere Teile hat.
Eine Aggregation wird wie eine Assoziation als Linie zwischen zwei Klassen dargestellt und zusätzlich mit einer Raute versehen. Die Raute steht auf der Seite des Aggregats, also des Ganzen. Die Komposition wird wie die Aggregation dargestellt, die Raute ist jedoch gefüllt. Assoziations- und Kompositionsbeziehungen können mit einer Multiplizitätsangabe, einem Beziehungsnamen und mit Rollenbezeichnungen dargestellt werden.	Ganzes 0..1 * Teil 1 1..* existenz-abhängiger Teil Abteilung Mitarbeiter

Beispiel: Klassendiagramm Hochschule



Quelle: UML-Benutzerhandbuch, S.125

Detailkonzept für Geoobjekte

Für den Entwurf von Geoobjekten gibt es keine standardisierte Vorgehensweise. Es wird hier deshalb eine Strategie aufgezeigt, mit der es möglich ist, Geoobjekte mit ihren relevanten geometrischen und semantischen Merkmalen und Beziehungen so zu beschreiben, dass sie sich mit verfügbaren GIS-Produkten realisieren lassen.

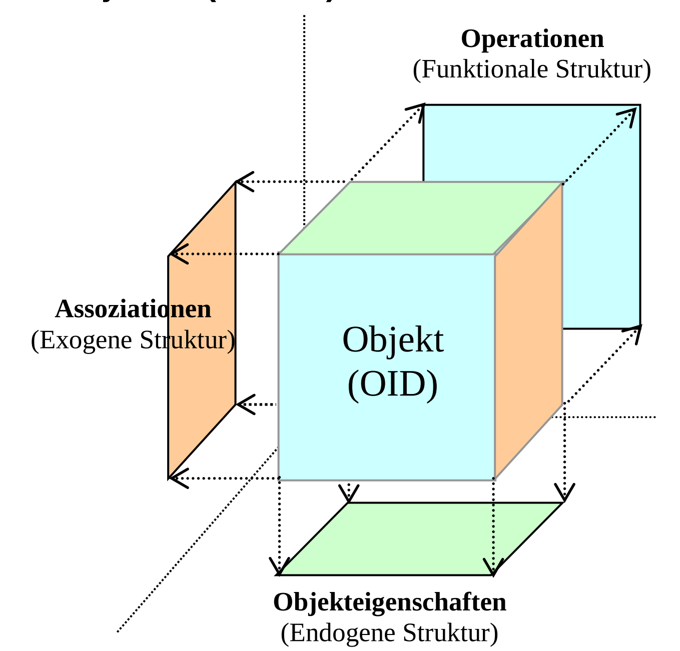
Der Detailentwurf für "allgemeine" Objekte (Klassen) kann in mehreren Teilsichten erfolgen, die Projektionen einer Entität unter verschiedenen Blickwinkeln entsprechen.

Drei Teilsichten für das Design von Datenobjekten (→ UML):

I. **Endogene** (innere) Struktur für die deskriptiven Merkmale in Form klassifizierter Objekteigenschaften (Attribute)

II. **Exogene** (äußere) Struktur für die Beziehungen (Assoziationen) zwischen Instanzen der eigenen Klasse und fremden Klassen

III. **Funktionale** Struktur für die Verhaltensmerkmale in Form von Methoden für den Zugriff auf das Objekt und seine Eigenschaften sowie Operationen für ihre Verarbeitung und Auswertung (Methoden)



Konfiguration von Geoobjekten

Nach *Bill/Fritsch* führt die vollständige Beschreibung der Merkmale von Geoobjekten zu einer **Dreiebenen-Hierarchie**, wobei die

- **unterste** Ebene die **Metrik** enthält, i.d.R. durch Koordinaten gegeben ist
- **mittlere** Ebene die **Topologie** aufnimmt
- **oberste** Ebene die **thematische** Bedeutung eines Geoobjekts wiedergibt.

Die metrische Information mit dem Geokode beruht auf der Geometrie des Raumes, die topologische auf der Geometrie der relativen Lage.

Für die drei Modellierungsebenen von Geoobjekten (Synonym: räumliches Objekt) werden folgenden Definitionen angegeben:

- Unter **geometrischem Modellieren** versteht man die Beschreibung, Bearbeitung und Speicherung der zugrunde liegenden Geometrie von räumlichen Objekten, indem analytische und approximierende Verfahren eingesetzt werden.
- Unter **topologischem Modellieren** versteht man die Beschreibung, Bearbeitung und Speicherung der Geometrie der relativen Lage von räumlichen Objekten. Als Hilfsmittel dienen dabei topologische Beziehungen und Invarianten.
- Unter **thematischem Modellieren** versteht man die Beschreibung, Bearbeitung und Speicherung der zugrunde liegenden Thematik eines räumlichen Objekts. Als Hilfsmittel dienen thematische Ebenen und Objekthierarchien, in denen verschiedene thematische Inhalte vorgehalten und Objekte zusammengefasst werden können.

Synthese der Teilmodelle

Die Vereinigung der drei Teilmodelle für die Geometrie, Topologie und Thematik kann man als Konfiguration der Geoobjekte betrachten.

Geoobjekt ::= < OID, Klasse, {Geometrieelemente}, {Attribute} >

Die Konfiguration legt die geometrischen Eigenschaften, die deskriptiven Merkmale (selbstbezogene Eigenschaften) und die Beziehungen (fremdbezogene Eigenschaften) der Geoobjekte fest und beschreibt sie mittels Geometrieelementen (spatial Feature) und Attributen.

Einige Autoren geben vier Modellierungsebenen an, als vierte Ebene kommt die Dynamik, d.h. die zeitliche Veränderung von Geoobjekten hinzu, nach *Streit* gilt:

- Die **Geometrie** umfasst alle Angaben zur räumlichen Lage und Ausdehnung des Geoobjekts auf Basis eines räumlichen Bezugssystems
- Die **Topologie** charakterisiert die räumlichen Beziehungen (Nachbarschaft) von Geoobjekten zueinander
- Die **Thematik** umfasst alle fachlichen Charakteristika einer Klasse von Geoobjekten
- Die **Dynamik** charakterisiert die zeitlichen Veränderungen der Geoobjekte.
  Diese können sowohl die Geometrie, die Topologie wie auch die Thematik der Geoobjekte betreffen

Die Modellierung zeitlicher Veränderungen in GIS wird als Spezialfall gesondert betrachtet, siehe temporale GIS (TGIS).

Aus der Kombination der drei Teilsichten für den konzeptionellen Entwurf von Objekten mit den drei Modellierungsebenen für die Konfiguration von Geoobjekten ergeben sich **fünf Teilentwürfe („Dimensionen“) für das Design eines Geoobjekts**

- **Thematisches Modell der inneren (endogenen) Struktur →Attribute**
- **Semantische Assoziationen für Beziehungen im thematischen Modell, äußere (exogene) Struktur →Attribute**
- **Geometrisches Modell der inneren (endogenen) Struktur →Geometrieelemente**
- **Toplogische Assoziationen für Beziehungen im geometrischen Modell, äußere (exogene) Struktur →Geometrieelemente oder separate Feature**
- **Operationen entsprechend den Verhaltensmerkmalen**

Die geometrischen und thematischen Eigenschaften bilden die deskriptiven Merkmale der Geoobjekte,
sie werden zur inneren Struktur (selbstbezogene) zusammengefasst

Die topologischen und semantischen Assoziationen beschreiben die Beziehungen der Geoobjekte,
sie bilden die äußere Struktur (fremdbezogene) der Geoobjekte

Objekt- und Beziehungstypen im Geodatenmodell

Objekttypen (Terminologie nach ATKIS)

REO: Raumbezogenes Elementarobjekt

- Elementarobjekte einer Klasse (Basisklasse)
- **REO ::= <OID, Klasse, {Geometrie}, {Attribut}> allgemein**
  Mindestens 1 Geometrieelement und mindestens 1 Attribut
- **REO ::= <OID, Klasse, Geometrie , {Attribut}>** ATKIS
  1 Geometrieelement und mindestens 1 Attribut
- Sind nach Form, Lage geometrisch und durch Attribute thematisch beschrieben
- Single- oder Multipart-Geometrie (siehe Simple Feature, Spatial Schema)

ZUSO: Zusammengesetzte Objekte

- Komplexe Objekte (Aggregate einer Basisklasse)
- Keine eigene Geometrie
- **ZUSO ::= <OID, Klasse, {REO},{Attribut}, [x,y]>**
- Mindestens 1 REO, zusätzliche Attribute
- ggf. Koordinate eines Bezugspunktes (z.B. Schwerpunkt) für räumliche Auswertungen

NREO: Nichtraumbezogenes Elementarobjekt

- Objekte ohne Raumbezug und Geometrie, die über Attribute mit Geoobjekten verknüpft werden können
- **NREO ::= <OID, Klasse, {Attribut}>**

Beziehungstypen zwischen Objekten in GIS

semantische Assoziation über Attribute, „vom Kleinen zum Großen" (thematisch)

- REO → ZUSO; über Attribute, Aggregation, Komposition
- ZUSO → NREO: über Attribute
- **statisch**, können im Modell abgebildet werden (Fremdschlüssel)
- Wichtige Daten für semantische Operationen sind georefrenzierte Adressen, Flurstücksnummern, Baublocknummern, Postleitzahlbezirke.

topologische Assoziationen über Inzidenzen (geometrisch)

- Zwischen REO-Instanzen derselben Klasse, die über gemeinsame Knoten/Kanten in Beziehung stehen (touches), d.h. räumlich zusammenhängen (Mosaik) und einen begrenzten Bereioch vollständig abdecken (Eulercharakteristik) – z.B. Flurstücke
- Haben keine Bedeutung, fass die REO-Instanzen nicht räumlich zusammnehängen (disjoint) – z.B. NSG, FFH, …
- **statisch**, können im Modell abgebildet werden (Inzidenzen)

topologische Relationen über den Raumbezug

- Ergeben sich erst dann, wenn die Instanzen zweier Klassen (Themen) aufgrund ihres einheitlichen Raumbezugs in Beziehung gesetzt (überlagert) werden – spatial relation
- Wichtige Relationen: within, contains – overlaps, intersects
- **nicht statisch**, können nicht statisch abgebildet werden, sondern müssen als Operatoren realisiert werden (siehe Egenhofer-Operatoren)

Teilentwürfe für das Geodatenmodell

Als Notationen für semantische Datenmodelle wird heute meist das UML-Klassendiagramm verwendet. Die UML erlaubt die Darstellung des semantischen Datenmodells in Form von Objekten und deren Beziehungen, eine Beschreibung, die dem Vorstellungsbild der Fachanwender gut entspricht. Die Darstellungselemente für das UML-Klassendiagramm (s.o.) sind jedoch nicht auf die speziellen Belange von Geoinformationen ausgerichtet, deren primäre Eigenschaften der Raumbezug, die Form und die geometrischen Beziehungen der räumlichen Entitäten sind.

Für die gute Lesbarkeit von semantischen Geodatenmodellen werden hier modifizierte Notationen für das UML-Klassendiagramm angegeben, mit denen die Besonderheiten von Geodaten stärker hervorgehoben werden können. Ziel ist es, dass eine Darstellung auf Grundlage von UML entsteht, in der einerseits der Anwender alle relevanten Entitäten seiner Welt, mit ihren für seine Anwendungen wesentlichen Eigenschaften und Beziehungen erkennen kann und andererseits für den Geoinformatiker eine eindeutige konzeptionelle Grundlage (Bauplan) für die Ableitung des logischen Entwurfs der Geodatenbank zur Verfügung steht. Das semantische Datenmodell ist eine wichtige Kommunikationsebene zwischen Anwender und IT-Spezialisten. Es soll das Datenmodell in einer für beide Seiten verständlichen, gut lesbaren Form, aus fachlicher Sicht, beschreiben.

Thematisches Modell und semantische Assoziation

Die Darstellung der thematischen Merkmale und der semantischen Assoziationen mittels Attributen erfolgt entsprechend den Vorgaben im UML-Standard (s.o.).

Attribute

Attribute bzw. Eigenschaften charakterisieren die thematischen Merkmale der Entitäten. Attribute besitzen einen Namen, einen Datentyp und entsprechende Werte, die Elemente eines zugeordneten Wertebereichs, ihrer Domäne sind.

Attribut ::= < Name, Datentyp, Wert (Domäne) >

Für multimediale Beschreibungen von räumlichen Entitäten können Attribute auch Verweise enthalten – Hyper Links – auf ergänzende Bild-, Ton-, Textdokumente oder Web-Seiten. Solche Medien können in Dateien oder als Binary-Large-Objects (BLOB) in Datenbanken gespeichert werden. Für umfangreiche verbale Beschreibungen können zukünftig nicht nur unstrukturierte Textdokumente (dokumentzentriert) sondern auch native XML-Strukturen (datenzentriert) verwendet werden. Die XML-Struktur erlaubt, dass der Inhalt der Dokumente in Suchfragen (Selektionen) einbezogen werden kann, d.h. transparent für die Datenbank ist.

Beziehungstypen für Geodaten: Assoziation, Aggregation, Komposition

Die semantischen Assoziationen entsprechen den Beziehungen mit Instanzen anderer Klassen auf Ebene der thematischen Merkmale. Dies sind insbesondere die Beziehungen zu Objekten, die primär keinen Raumbezug haben (z.B.: Flurstück → Eigentümer; Haus → Einwohner) oder Aggregation untergeordneter Objekte darstellen.

Bei strengen „ist Teil von“-Beziehungen (Kompositionen) ist zu beachten, dass das Löschen des Ganzen auch zum Löschen der Teile führen sollte. Zwischen verschie denen Klassen können zudem Vererbungsbeziehungen (Generalisierung, Spezialisierung) bestehen, dies sind Klassen, die miteinander in einer enger Beziehung stehen.

Geometrisches Modell (2D)

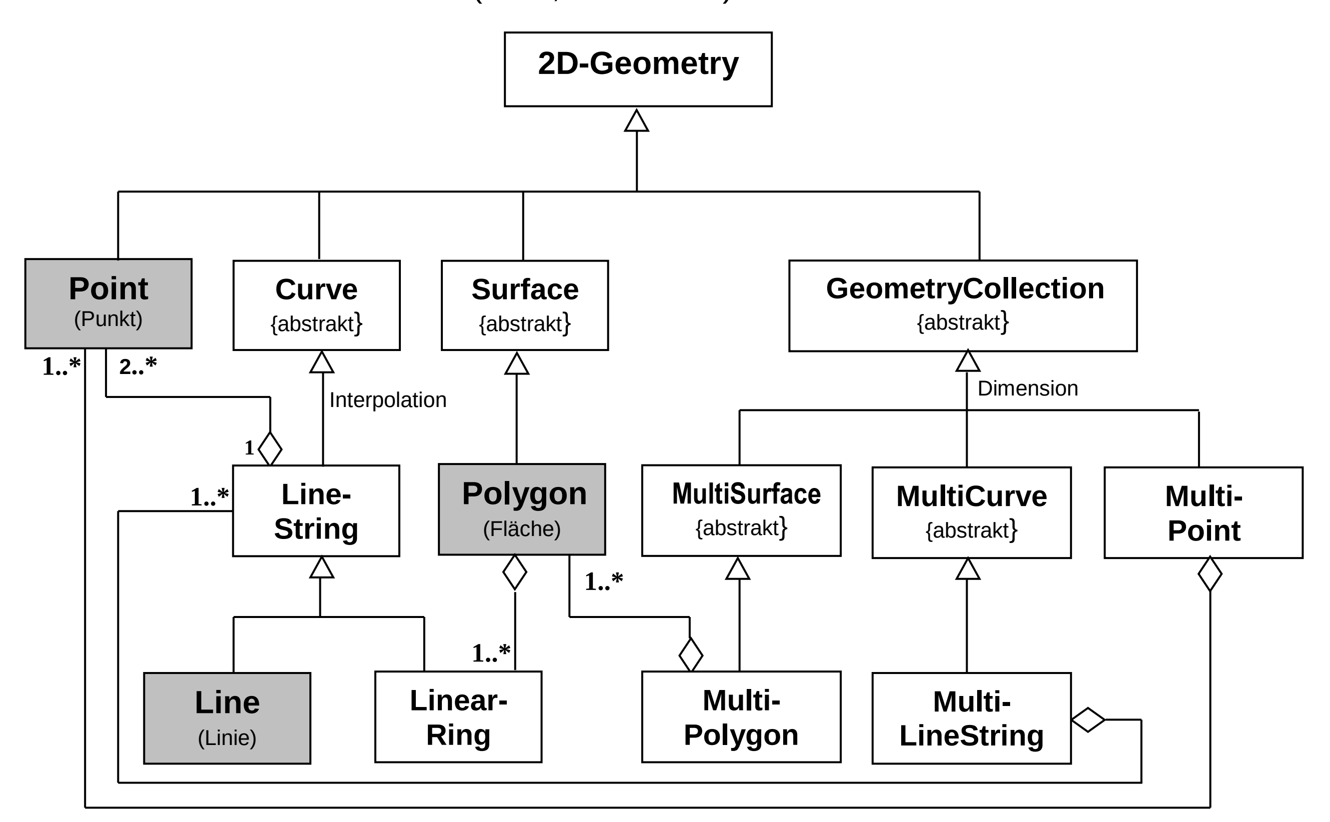
Die Geometrie umfasst alle Angaben zur absoluten räumlichen Lage, Ausdehnung und Form eines Geoobjekts auf Basis eines festgelegten räumlichen Bezugssystems (Koordinatensystem). Für die Randbeschreibung und Geocodierung einer Entität reichen im Prinzip die einfachen Geometrieelemente: **Punkt, Linie, Fläche** (feature type) aus.

Die Entitäten werden approximativ durch diskrete Randpunkte beschrieben. Punkte werden durch gerade oder gekrümmte Linien verbunden; geschlossene Linienzüge (Polygone) bilden die Konturen von Flächen. In der Praxis hat sich jedoch gezeigt, dass zusätzlich multiple und komplexe Grundformen – als Bauteile – für die Modellierung von Geoobjekten von Vorteil sind, die durch Aggregation der einfachen Elemente entstehen.

Konzeptionelle Schemata für Geometrieelemente enthalten die Normen:

- ISO 19125: Simple Features (einfach) → *GML v2 (= Geography Markup Language)*
- ISO 19107: Spatial Schema (mit Topologie) → *GML v3*

Simple Feature Spezifikation beschreibt ein herstellerunabhängiges 2D-Geometriemodell. Die Elemente werden als „simple“ (einfach) bezeichnet, wenn sie keine Schlingen (Selbstüberschneidungen) oder tangentiale Selbstberührungen aufweisen. Die Topologie wird nicht abgebildet. Die Simple Feature Spezifikation ist implementiert z.B. in der Datenbank ORACLE, der Programmiersprache Java und dem normierten Datenaustauschformat *GML v2* (GML, ISO 19136).



Die geometrischen Grundformen (feature type) sind **Point, Line und Polygon** (singlepart feature). Daraus können durch Aggregation multiple Formen (multipart feature) gebildet werden. Als Linien sind nur Geraden vorgesehen. Die abstrakte Klasse GeometryCollection definiert eine Menge gleichartiger, einfacher oder multipler Elemente (feature class).

Spatial Schema definiert, gegenüber Simple Feature, ein erweitertes konzeptionelles Basisschema für die geometrischen und topologischen Eigenschaften von Geodaten. Die Geometrie und Topologie werden als Attribute der Objekte behandelt. Es gibt die geometrischen Grundformen: Punkt, Linie, Fläche, Körper und die topologischen Grundformen: Knoten, Kante, Masche, Raumelement.

- Spatial Schema unterstützt Vektordaten in bis zu 3 Dimensionen
- Räumliche Charakteristiken werden als räumliche Attribute dargestellt
- Die Abbildung der Topologie ist mit eigenen Features ist möglich
- Es gibt auch gekrümmte Linien (Interpolation: Arc, Spline, Klothoide, ...)
- Spatial Schema definiert topologische Operationen („Egenhofer Operatoren“)

Das Spatial Schema ist Grundlage für den standardisierten Datenaustausch mit *GML v3* und soll zukünftig auch Grundlage für moderne Geodatenmodelle in den GIS-Produkten werden.

Für die Darstellung im semantischen Datenmodell reicht die Angabe des Elementtyps aus, für die genaue Spezifizierung wird bei der logischen Modellierung Bezug auf die Normen bzw. auf die Besonderheiten der verfügbaren Datenstrukturen im jeweiligen GIS-Produkt genommen, die heute noch sehr unterschiedlich sind.

Geometrieelement spatial feature Darstellung wie Attribut Notation und Beispiel	Ein Geometrieelement beschreibt ein räumliches Merkmal (feature) eines Geoobjektes. Es ist in jedem Geoobjekt einer Klasse gleichermaßen enthalten und wird von jedem Geoobjekt mit einer individuellen Lage und Form repräsentiert. Geometrieelemente haben keine Identität, sie sind carakteristischeTeile (feature) von Geoobjekten.
Geometrieelemente werden mindestens mit ihrem Geometrietyp aufgeführt und können zusätzliche Angaben zu einem Initialwert, Eigenschaftswerte und Zusicherungen enthalten. Der Typ beginnt mit einem Kleinbuchstaben, Eigenschaftswerte und Zusicherungen stehen in geschweiften Klammern. Sie werden in der Klasse vor die Attribute geschrieben, da sie den geometrischen Typ eines Objektes angeben. Die Rubriken → Klassenname, Geometrieelemente, Attribute, Operationen – werden durch horizontale Linien getrennt.	Sichtbarkeit feature type: Initialwert {Zusicherung} ↑ (+ public, # protected, - private) + sPoint: (10,10) {x>0,y>0} + sLine: (interpolation="spline")

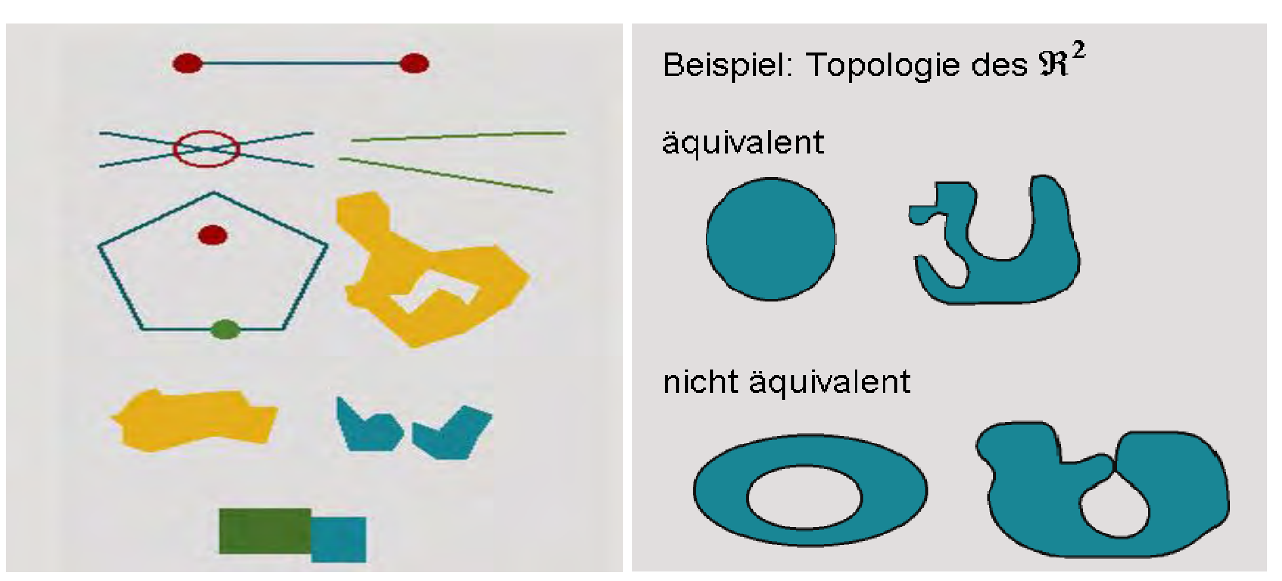
(Mögliche) Bezeichnungen für die Vektor-Geometrieelemente

Feature Type		Bezeichnung	Dim	Bemerkung
Single-Part	Punkt	sPoint	0D	Punktobjekt oder topolog. Knoten
	Linie	sLine interpolation = "…"	1D	linear (Gerade), arc (Bogen), spline, klothoide, ...
	Fläche	sPolygon	2D	geschlossene Struktur aus Linien
Multi-Part Elemente		mPoint, mLine, mPolygon		Mehrfach-Punkt, -Linie, -Fläche hängen räumlich nicht zusammen
Komplexe Geometrie		kGeometry	1D 2D	zusammengesetzte Geometrie aus Elementen mit verschiedenem Typ

Toplogische Assoziationen und Relationen

Räumlichen Beziehungen (Nachbarschaften) von Geoobjekten können unabhängig von einer Metrik durch ihre Topologie charakterisiert werden, d.h. unabhängig von Winkel-/ Streckenmessungen und invariant gegenüber Koordinatentransformationen. Die Topologie wird auch als die Geometrie der relativen Lage oder „Gummituchgeometrie" bezeichnet und beruht auf den topologischen Invarianten und den Beziehungen zwischen Knoten, Kanten und Maschen (Inzidenzen).

Toplogischen Invarianten im zweidimensionalen Raum sind die Geschlossenheit, Schnittpunktreue, Trennung innen/außen und die Randpunkteigenschaft.



Auf der Grundlage der topologischen Invarianten lassen sich räumliche Relationen zwischen unabhängigen Geoobjekten definieren, die nicht im Datenmodell (statisch) abgebildet werden können, da sie i.d.R. erst aus Überlagerungen unabhängiger Objekt-bereiche resultieren. Für die Auswahl von Geoobjekten nach räumliche Relationen sind deshalb Operatoren erforderlich, die auf Ebene einer Metrik die gegenseitige Lage der Geoobjekte (z.B. Gebäude ↔ Flurstück) entsprechend einer Relation prüfen. Beispiele für räumliche Relationen sind: „innerhalb", „schneiden sich" (siehe Kapitel 9).

Topologische Beziehungen zwischen den Knoten, Kanten und Maschen einer Geometrie werden als

- **Inzidenz** bezeichnet, wenn sie verschiedene Elemente und
- **Adjazenz** bezeichnet, wenn sie gleichartige Elemente verbinden.

Es gibt vier topologische Grundbeziehungen zwischen Knoten, Kanten und Maschen:

- **Knoten-Adjazenz:** Zwei Knoten sind über eine Kante miteinander verbunden (benachbart)
- **Knoten-Kanten-Inzidenz:** Alle Kanten, die von einem Knoten abgehen sind mit ihm inzident
- **Maschen-Adjazenz:** Zwei Maschen sind über eine Kante miteinander verbunden (benachbart)
- **Maschen-Kanten-Inzidenz:** Alle Kanten einer Masche (Umringskontur) sind mit ihr inzident

Aus diesen vier Grundbeziehungen können weitere Beziehungen abgeleitet werden, wie z.B. die Knoten-Maschen-Inzidenz.

Inzidenzbeziehungen setzen voraus, dass Geoobjekte gemeinsame Knoten und Kanten haben, über die sie miteinander verknüpft sind. Dies sind i.d.R. Geoobjekte derselben Klasse (z.B. Flurstück-Flurstück) oder solche, die auch in der Natur gemeinsame Grenzen oder Punkte haben. Die Inzidenzen können als statische Merkmale im Datenmodell abgebildet werden und entsprechen den äußeren Beziehungen der Geoobjekte, ihren toplogischen Assoziationen. Inzidenzbeziehungen sind immanente Eigenschaften von Geoobjekten, wie thematische und metrische Merkmale.

Bei der zweidimensionalen Randbeschreibung von Geoobjekten besteht ein einfacher Zusammenhang zwischen metrischer und topologischer Information. Die Topologie kann aus jeder konkreten Metrik abgeleitet werden. Ergeben sich die topologischen Beziehungen aus der Metrik eines Raumes, bezeichnet man sie auch als

induzierte Topologie, wobei die
Punkte → Knoten, Linien → Kanten, Flächen → Maschen induzieren.

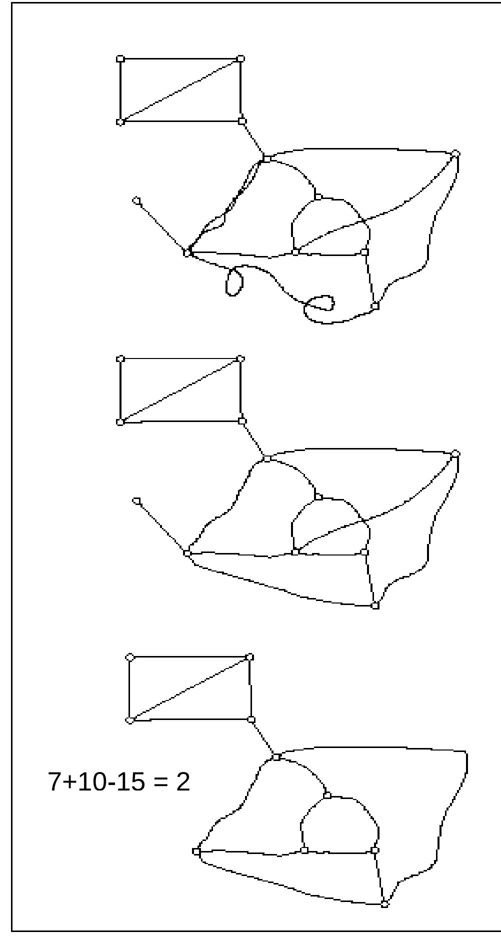
Die induzierten topologischen Beziehungen zwischen Knoten, Kanten und Maschen ergeben die exogene geometrische Struktur mit den äußeren Beziehungen der Objekte.

Darstellung topologischer Beziehungen

Die Topologie beschreibt die Beziehungen zwischen Knoten und Kanten in einer Geometrie, unabhängig von einer konkreten Metrik. Ein **Graph** ist allgemein eine Darstellung von Elementen und deren Beziehungen durch Knoten und Kanten. Wegen der engen Verwandtschaft der beiden Konzepte, ist es naheliegend für die Abbildung topologischer Beziehungen Graphen zu verwenden. Nachfolgend werden Graphen sowohl für die Abbildung der durch die Geometrie induzierten Topologie (i.d.R. planare Struktur) als auch für die Darstellung der Beziehungen zwischen den Geometrie-elementen verwendet.

Sätze für planare, ungerichtete Flächengraphen (4-FarbenTheorem, Landkarte):

- Ein **endlicher Graph** G={N,K} ist durch die begrenzte Menge seiner Knoten und Kanten definiert. Als Knoten (N = Node) werden die Stellen bezeichnet, in der eine Kante (K) beginnt oder endet, oder in der sich mehrere Kanten treffen. Eine Kante ist die Verbindung zweier Knoten, jede Kante hat einen Anfangs- und einen Endknoten.
- Ein **einfacher Graph** ist ein endlicher Graph ohne Mehrfachkanten und ohne Schlingen.
- Ein **einfacher Graph heißt zusammenhängend**, wenn es zu je zwei beliebigen Knoten einen Weg gibt, der die Knoten verbindet. Eine Verbindung heißt Weg, wenn kein Knoten zwischen Anfangs- und Endknoten mehrfach durchlaufen wird.
- Ein **einfacher Graph heißt planar**, wenn er in der Ebene so dargestellt werden kann, dass keine Schnittpunkte von Kanten vorkommen, die nicht zugleich Knoten sind.
- Ist G={N,K} ein **zusammenhängender planarer Graph**, so zerlegen die Kanten die Ebene in endlich viele zusammenhängende Gebiete, die als Maschen (M) bezeichnet werden, von denen genau eine, das äußere Gebiet, der Außenraum, nicht beschränkt ist. Für zusammenhängende planare Graphen gilt die Euler-Charakteristik: **M + N - K = 2**, falls der unbegrenzte Außenraum als eine Masche mitgezählt wird.
- Die Kanten einer Masche bilden einen Weg, der als einfacher Zyklus oder topologischer Kreis bezeichnet wird. **Einfache Zyklen** sind Wege in Graphen, die von einem Knoten ausgehen und im selben Knoten enden, wobei die Anfangs- und Endkanten voneinander verschieden sind.
- Ein **ebener Flächengraph** oder Landkarte L= {N,K,M} ist ein planarer Graph G={N,K} mit den zugehörigen Ländern M. Jede Kante trennt zwei verschiedene Maschen, jede Masche ist von einem einfachen Zyklus begrenzt. Jeder Knoten hat mindestens 2 inzidente Kanten.
- Für ebene Flächengraphen L= {N,K,M} aus k Kontinenten (nicht zusammenhängende Teile) gilt: **M + N – K = k + 1**. Die Euler-Charakteristik ist eine notwendige, jedoch nicht hinreichende Bedingung für fehlerfreie, planare Flächenstrukturen.



Endlicher Graph G={**N,K**}
Ein Graph ist eindeutig durch die Angabe seiner Knoten- (**N**) und Kantenmenge (**K**) sowie durch die Inzidenzabbildung der Knoten auf die Kanten definiert.

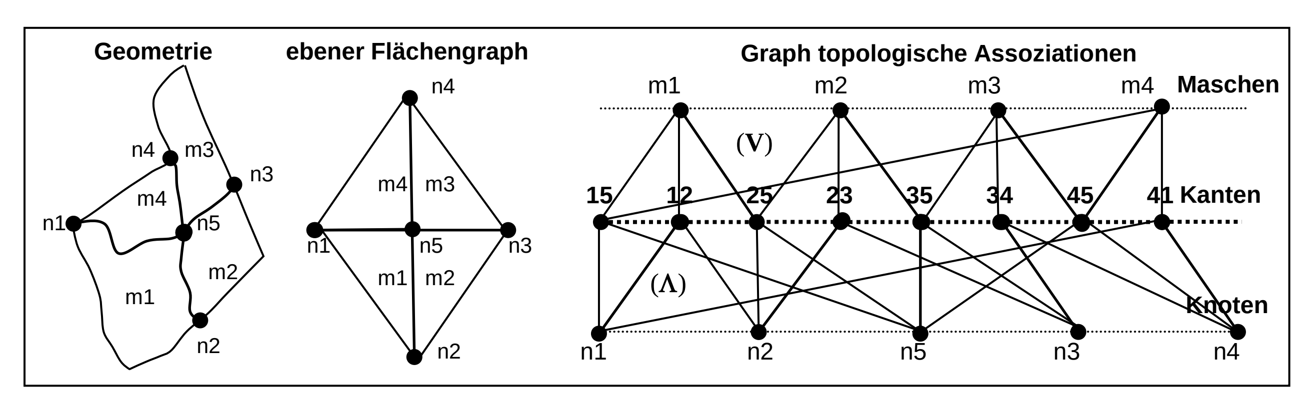
Einfacher Graph hat keine Mehrfachkanten und keine Kanten mit Schlingen („simple feature").
Er heißt **zusammenhängender Graph,** wenn zwei beliebige Knoten durch einen Weg miteinander verbunden sind.

Planarer Graph
Es gibt keine Schnittpunkte von Kanten, die nicht zugleich Knoten sind
Die Kanten eines zusammenhängenden planaren Graphen zerlegen die Ebene in endlich viele Maschen (**M**), von denen genau eine, das äußere Gebiet (Außenraum), nicht beschränkt ist.

Ebener Flächengraph (Landkarte) **L**={**N,K,M**}
Ist ein planarer Graph mit den dazugehörenden Maschen (Länder), die durch einfache Zyklen begrenzt sind. Ebene Flächengraphen erfüllen die Euler-Charakteristik in der Form: **M + N - K = 2**, wobei der Außenraum als Masche mitgezählt wird

Darstellung topologischer Beziehungen mit Graphen

Betrachtet man die Strukturen einer zweidimensionalen Flächengeometrie in den Sichten: Geometrie, induzierte Topologie und Beziehungsstruktur, so zeigt sich, dass die Beziehungen zwischen den Elementen des Graphen und damit auch zwischen den Geometrieelementen ein Netzwerk bilden!

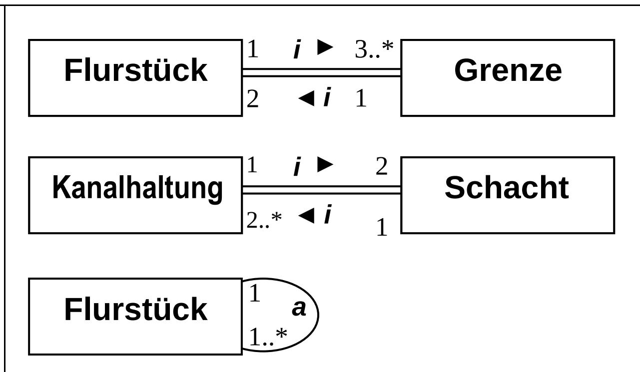


Aus der Darstellung der topologischen Assoziationen (Beziehungen) lassen sich die grundlegenden Inzidenzen und Adjazenzen direkt erkennen. Inzidenzen entsprechen den Verbindungen zweier verschiedenartiger Elemente. Adjazenzen zeigen sich als „(**Λ**)-förmige" Struktur für die Knoten und als „(**V**)-förmige" Struktur für die Maschen, ausgehend von einer Kante, sie implizieren immer zwei Inzidenzen. Zwei Knoten oder zwei Maschen sind also immer über eine Kante miteinander adjazent, d.h. verbunden. Bei Randkanten bildet der Außenraum die zweite Masche. Kanten sind damit bei einer Randbeschreibung von Flächen das „Adjazenzmedium" für Knoten und Maschen.

Topologische Assoziation

Topologischen Beziehungen auf Grund gemeinsamer Knoten und Kanten von Geoobjekten, den Inzidenzen oder Adjazenzen, werden hier als topologische Assoziation *(Duden: Assoziation = „ursächliche Verknüpfung“)* bezeichnet.

Sie können im Datenbestand statisch abgebildet werden.

Topologische Assoziation spatial association Notation und Beispiel	Eine topologische Assoziation beschreibt eine geometrische Beziehung zwischen den Geoobjekten derselben Klasse oder Geoobjekten verschiedener Klassen auf Ebene der Inzidenzbeziehungen gemeinsamer Knoten und Kanten
Eine topologische Assoziation wird durch eine Linie zwischen den beteiligten Klassen dargestellt. An die Linie wird der Assoziationstyp kursiv geschrieben. Damit die Klassen und die Beziehung in richtiger Reihenfolge gelesen werden, kann neben den Beziehungstyp ein kleines ausgefülltes Dreieck in Leserichtung gezeichnet werden. An den jeweiligen Enden kann die Multiplizität der Beziehung angeschrieben werden. Zulässige Assoziationstypen: i(nzident), a(djazent) Adjazente Assoziation nennt man refelxiv, da sie zwischen Objekten der selben Klasse besteht.	

Topologische Relation

Topologische Beziehungen zwischen Geoobjekten, die auf den topologischen Invarianten beruhen, lassen sich mit Hilfe des auf Egenhofer zurückgehenden Dimensionally Extended 9-Intersection Model (DE-9IM), das auch als Neuner-Verschneidungs-Matrix (NVM) bezeichnet wird, beschreiben. Sie werden hier als topologische Relation *(Duden: Relation = „Beziehung, in der sich 2 Dinge vergleichen lassen“)* bezeichnet.

Sie können nicht statisch im Datenbestand abgebildet werden, sondern sind bei Bedarf durch geeignete Operationen zu realisieren (siehe Kapitel 7). Die **Operatoren** vergleichen die Geometrien zweier Themen bezüglich eines gewählten Kriteriums und liefern die Menge der Geoobjekte zurück, die das Kriterium erfüllen.

Topologische Relation spatial relationship Notation und Beispiel	Eine topologische Relation beschreibt eine geometrische Beziehung zwischen den Geoobjekten auf Ebene topologischer Invarianten (DE-9IM, Egenhofer).
Eine topologische Relation wird durch eine gestrichelte Linie zwischen den beteiligten Klassen dargestellt. An die Linie wird der Stereotyp (Schlüsselwort) für die <<relation>> geschrieben. Durch einen Pfeil kann die Relation gerichtet werden. Damit die Klassen und die Beziehung in richtiger Reihenfolge gelesen werden, kann die Richtung der Beziehung durch Pfeile an den Linienenden gezeichnet werden. An den jeweiligen Enden kann die Multiplizität der Beziehung angeschrieben werden. Zulässige Relationstypen: entsprechend den Egenhofer-Operatoren: contains, within, touches, disjoint, equals, crosses overlaps, intersects	Flurstück 1 - - - - - - -> * Wohnhaus <<contains>> Bestand * <- - - - - - -> * Standort <<intersects>>

Anmerkung: Topologische Relationen sind im semantischen Datenmodell nur dann darzustellen, wenn diese Beziehung eine Integritätsbedingung zwischen verschiedenen Themen darstellt. Die Prüfung der Bedingung kann z.B. über Datenbanktrigger erfolgen. Allgemein ergeben sich topologische Relationen i.d.R. erst bei der Überlagerung verschiedener Themen.

Topologie im Geodatenmodell

für exogene Beziehungen auf Ebene der Geometrie

Topologische Assoziation	**Topologische Relation**
Assoziation = DUDEN: ursächliche **Verknüpfung**	**Relation** = DUDEN: **Beziehung** zwischen vergleichbaren Dingen
Inzidenzen (Knoten-Kanten-Maschen)	**Topologische Invarianten**
BeziehungsGraph	**Geometrie**, bzw. Graph der induzierten Topologie
statisch → speicherbar (persistent) → **Datenmodell** → geeignete Datenstrukturen → separate / integrierte Speicherung	**dynamisch** → entstehen erst bei der Überlagerung von verschiedenen Objektklassen → **Operatoren** erforderlich, in DB ggf. über Trigger[1] realisieren
Vermeidung von Redundanzen Topologische Navigation	Auswahl (Selektion) von Objekten nach räumlichen Beziehungen
Eulercharakteristik M+N-K=2 → notwendige, aber nicht hinreichende Bedingung in ebenen Flächengraphen M= Maschen (face) N = Knoten (node) K = Kanten (edge)	**Operatoren** (Egenhofer) *DE-9IM* heißt *Dimensionally Extended 9-Intersection Model* vollständig innen: **contains**, **within**, equals vollständig außen: disjoint, touch schneiden sich: **overlaps**, **intersects,** crosses
Beispiel: Objektklasse Flurstück Die Instanzen decken das Projektgebiet vollständig ab. Sie sind über gemeinsame Knoten und Kanten verknüpft. Die Flurstücke bilden adjazente Maschen, die über gemeinsame Grenzen (Kanten) direkt benachbart sind.	Beispiel: Objektklassen Flurstück, Gebäude Durch die Überlagerung der beiden Klassen Flurstück und Gebäude ergeben sich topologische (räumliche) Beziehungen. z.B. Flurstück contains Gebäude (enthält) z.B. Gebäude within Flurstück (innerhalb)
Beispiel: Objektklasse Gebäude Die Instanzen sind nicht direkt benachbart (adjazent), damit treten keine topologischen Assoziationen (Inzidenzbeziehungen) auf	Die topologischen Relationen sind aufgrund der Unschärfe der Geometrie (Genauigkeit der Metrik), die die Beziehungen induzieren, ebenfalls nur unscharf auszuwerten. Für Operatoren wie contains, within, intersects ergeben sich meist befriedigende Ergebnisse. Für Operatoren wie touch, disjoint, equals erlauben die GIS-Produkte die Angabe von Pufferbreien für die Kanten, so dass auch diese Operatoren sinnvolle Ergebnisse liefern. Die Pufferbreite muss in Abhängigkeit von der Genauigkeit der Kanten gewählt werden. Falls topologische Relationen im Datenbestand als Integritätsbedingungen formuliert werden sollen, können sie mittels Trigger[1] realisiert werden.

[1] Über Trigger lassen sich Anweisungen festlegen, die vom Datenbankserver automatisch gestartet werden, wenn ein bestimmtes Ereignis (before, after | delete, insert, update) für eine Tabelle eintritt. Sie können verwendet werden z.B. für die Automatische Prüfung der Datenkonsistenz oder die automatische Sicherung der semantischen Integrität.

Beispiel: Semantisches Datenmodell für Forstbetriebsdaten

Fachanforderungen (Zusammenstellung)

Pragmatik	**wichtige Merkmale:**
Abbildungsbereich	Land Bayern, Staatswald (850.000 ha)
Erfassungszeitraum	1993-2002 (Ersterfassung)
Erfassungsmethode	manuelle Digitalisierung, Vektordaten
Basisobjekt: Bestand	1 bis n Teilflächen, Basisbezugsraum für waldbauliche und Betriebswirtschaftliche Maßnahmen

Geometrie	**wichtige Eigenschaften**
Linien	hierarchisches System von Grenzen
Flächen	hierarchisches System von Flächen
Erfassungsmaßstab	1:5.000-1:10.000
Mittl. Datendichte (Geo)	1450 Elemente je Quadratkilometer
Digitalisierung	Ersterfassung nach Folien, Fortführung objektweise
Raumbezugssystem	GK-System (12°-Meridian)
Anschlusselemente	Blattecken der Flurkarte 1:5.000
Verteilung	Raster, Abstand ca.2.3 km, 14.033 Punkte, Genauigkeit ± 0.5 m
Anschluss	Helmert- oder Affintransformation mind. 4 homologe Punkte

Grenzlinien-Hierarchie	**Linienart**	**Puffer**
Staatswaldgrenze	(1) Flurstücksgrenze	± 2 m
Wege/Gewässer	(2) topogr. Linie / scharf	± 3 m
ständige forstw. Grenzen	(3) topogr. Linie / unscharf (Grate, Rinnen)	± 5 m
Bestandsgrenzen	(4) Vegetationsübergang	± 10 m

Flächen-Hierarchie	**Kontur aus Linienarten**
Forstamt (Staatswald)	(1)
Distrikt	(1) / (2)
Abteilung / Unterabteilung	(1) / (2) / (3)
Bestand	(1) / (2) / (3) nach außen, (4) gegeneinander

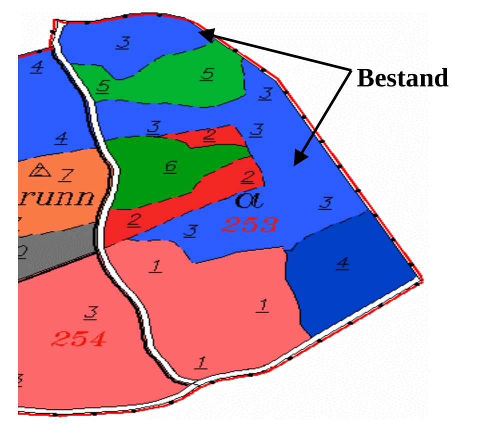
Detailbeschreibung des Basisobjektes Bestand

Auszug aus der Fachanforderung: *Die Forstbetriebskarte enthält die farbige Darstellung der Pflege- und Nutzungsarten oder der Entwicklungsstadien mit den geplanten Pflegemaßnahmen von (Wald)-Beständen. Der Bestand ist ein Kollektiv von Bäumen innerhalb eines abgegrenzten Gebiets, das eine einheitliche waldbauliche Bewirtschaftung ermöglicht.*

Er ist die kleinste Verwaltungs- und Bewirtschaftungseinheit im Staatswald *und bildet damit den Basisbezugsraum für die beschreibenden Daten der Forsteinrichtung sowie für andere wichtige Zustands und Planungsgrößen.*

Ein Bestand kann aus mehreren, ***räumlich nicht zusammenhängenden Teilflächen bestehen.*** *Organisatorisch werden Bestände der ständigen Waldeinteilung zugeordnet, die eine* ***Hierarchie*** *aus Bestand, Abteilung, Distrikt und Forstamt darstellt.*

Die waldbaulichen Eigenschaften des Bestands werden durch Standortsdaten ergänzt, die einer gesonderten Standortskartierung ausgewiesen werden. Standorte sind zusammengefasste Flächen, die hinsichtlich der ökologischen Voraussetzungen, den waldbaulichen Möglichkeiten und der Gefährdung des Baumbestandes ähnliche Voraussetzungen haben. Ihre Farbgebung richtet sich nach dem Substrattyp (Bodenart)...“.



Für den Bestand gibt es folgende Fachanforderungen, Beispiele sind: Farbige Darstellung in der Forstbetriebskarte, Flächenberechnung und Abgleichung auf die Katastersollfläche, thematische Ausprägungen für Attribute der FE/DB, z.B. Bestandsform, Verschneidung mit Standorten

Anforderungsanalyse

Die Grundlage für die Konzeption eines Datenmodells ist die sorgfältige Analyse der Fachanforderungen (externes Schema). Dazu sind Fragen zu klären wie z.B.:

- Welche Entitäten der realen Welt sind als Geoobjekte abzubilden ?
- Welche Merkmale haben diese Geoobjekte (geometrische, thematische) ?
- Welchen Beziehungen (Assoziationen) zwischen den Geoobjekten bestehen (semantische, toplogische) ?
- Welche Hierarchien gibt es ?
- Welche Operationen und Funktionen werden benötigt ?

Die Beantwortung dieser Fragen erfordert eine systematische Vorgehensweise zur Präzisierung der Problemstellung, um daraus die Entwicklung einer qualifizierten Spezifikation zu ermöglichen.

Das Ziel der Spezifikation ist es zu beschreiben, welche Leistungen ein System erbringen soll, aber nicht wie es das tun soll. Das Ergebnis soll eine vollständige, konsistente und eindeutige Anforderungsdefinition für das System sein.

Dieser Prozess wird auch als Requirements Engineering bezeichnet.

Das methodische Vorgehen bei der Analyse der Fachanforderungen kann man grob in folgende Schritte unterteilen:

- **Objekte und Klassen identifizieren**: Klassen und Objekte entsprechen oft Substantiven in den Fachanforderungen. Objekte und Klassen trennen → Objekte sind ansprechbar, Klassen sind anonym.
- **Attribute identifizieren**: Aus Sicht eines Objekts nach allgemeinen oder problemspezifischen Beschreibungen (Eigenschaften) suchen. Attribute haben keine Identität, sie sind Teile von Objekten.
- **Assoziationen identifizieren**: Beziehungen zwischen Klassen entsprechen oft Verben oder Verbalphrasen der Fachanforderung (keine Aktionen). Hierbei ist auch nach Aggregationen und Kompositionen zu suchen, die oft durch Ausdrücke wie „ist Teil von" oder „besteht aus" beschrieben werden.
- **Verfeinern durch Vererbung**: Spezialisierungen werden angezeigt, wenn eine Klasse alle Merkmale einer anderen sowie noch einige weitere Merkmale hat. Generalisierungen werden angezeigt, wenn zwei Klassen viele gleiche aber auch einige unterschiedliche Merkmale aufweisen, d.h. die Oberklasse fehlt.
- **Suche nach Operationen** (Dienste): Klassen separat untersuchen und mit Diensten aus der Anwendungswelt versehen, dann Hauptaufgaben des Systems untersuchen und prüfen, ob die notwendigen Operationen vorhanden sind.
- **Überarbeitung des Ergebnisses** (Reduktion): Klassen mit nur einem Objekt oder Objekte mit nur einem Dienst, Attribut oder Beziehung sowie abgeleitete Attribute oder Objektbeziehungen sind vermutlich überflüssig.

Literatur:

Partsch H. (1998) *Requirements-Engineering systematisch, Springer-Verlag, Berlin und Heidelberg*

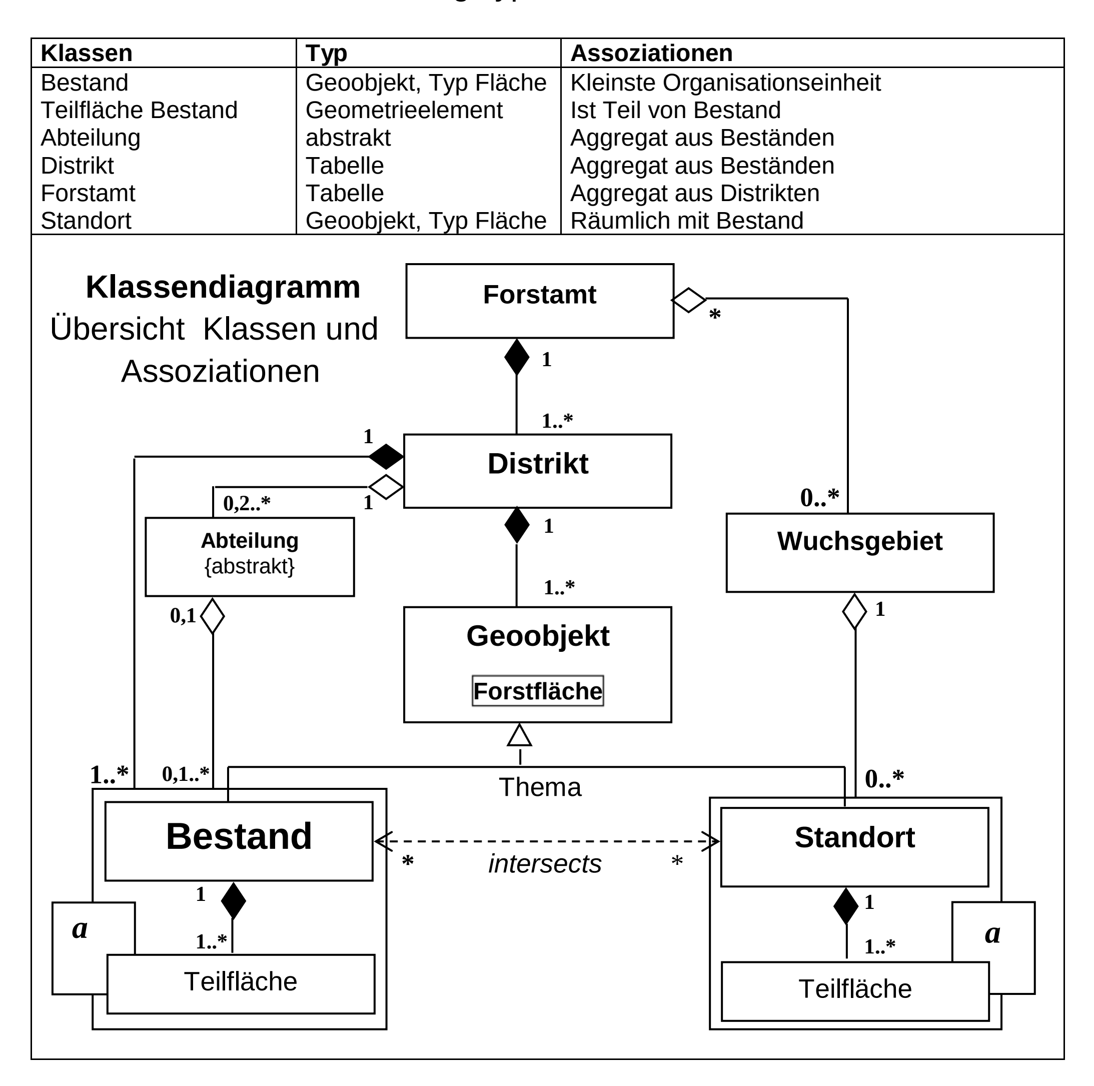
Nachfolgend werden die wesentlichen Aspekte bei der Konzeption von Geoobjekten am Beispiel des (Wald)Bestandes in Form eines Klassendiagramms aufgezeigt. Die Struktur des Bestandes ist repräsentativ für andere flächenhafte Geoobjekte.

Klassendiagramm

Der **konzeptionelle Entwurf** auf Grundlage der erstellten Anforderungsdefinition beschreibt das Datenmodell in einer formalen Darstellung als semantisches Datenmodell, unabhängig von seiner Realisierung mit dem speziellen Datenmodell eines konkreten GIS/Datenbank-Produktes.

Aus Systemsicht sind Datenbanken statische Strukturen aus Objekten (Datenobjekte) und deren Beziehungen sowie geeignetem Operationen für ihre Verwaltung, Auswertung und Präsentation.

Für die Darstellung der Klassen und ihrer Beziehungen wird im konzeptionellen Schema das Klassendiagramm verwendet, das eine Gesamtsicht der statischen Struktur wiedergibt. Das Klassendiagramm kann unter Anwendung der angegebenen Regeln für die Problemanalyse aus den Fachvorgaben abgeleitet werden. Es zeigt die Klassen, ihre Assoziationen und den Beziehungstyp.



Klassen	Typ	Assoziationen
Bestand	Geoobjekt, Typ Fläche	Kleinste Organisationseinheit
Teilfläche Bestand	Geometrieelement	Ist Teil von Bestand
Abteilung	abstrakt	Aggregat aus Beständen
Distrikt	Tabelle	Aggregat aus Beständen
Forstamt	Tabelle	Aggregat aus Distrikten
Standort	Geoobjekt, Typ Fläche	Räumlich mit Bestand

Die Teilflächen sind aus fachlicher Sicht keine Geoobjekte, sie werden deshalb als Teilobjekt Geometrie, basierend auf einem Flächenelement (Polygon) realisiert und über eine 1:n-Relation mittels attributiver Merkmale zum Geoobjekt verbunden. Die symbolische Multiplizität **a** (Adjazenz) bezeichnet dabei die topologischen Beziehungen zwischen den Geoobjekten über die Maschen-Adjazenz, sie bedeutet: 1 Teilfläche ist benachbart mit 0..n anderen Teilflächen, wobei die Spezialfälle 0 = isolierte Fläche und 1 = Randfläche bedeuten.

Neben den Beziehungen im Klassendiagramm sind der Raumbezug und die Geometrie die wesentlichen Designfaktoren für ein Geodatenmodell. Sie wirken sich auf die endogene und die exogene Struktur aller Geoobjekte aus.

Bei GIS, die einen weiten Maßstabsbereich – von der Detaildarstellung bis zur Übersichtsdarstellung – umfassen, ist es ggf. nötig für dasselbe Geoobjekt mehrere geometrische Repräsentationen mit verschiedenen Abstraktionsgraden vorzuhalten, da die Generalisierung heute nur in Ansätzen algorithmisch gelöst ist.

Das semantische Datenmodell (Zusammenfassung, nach Schilcher)

- Im Rahmen des konzeptionellen Entwurfs erfolgt eine Strukturierung der Daten unter Einsatz bestimmter Modellierungs- und Beschreibungstechniken (z.B. UML).
- Der Übergang von der realen Welt zum Modell erfolgt durch einen, mit einem bestimmten Detaillierungsgrad durchzuführenden Abstraktionsprozess.
- In dieser Phase wird festgelegt, welche Objekte der realen Welt im Einzelnen zu berücksichtigen sind, welche Eigenschaften der Objekte erfasst werden müssen und welche Beziehungen zwischen den Objekten vorhanden sind. Weiterhin werden gleichartige Objekte in Klassen zusammengefasst, um eine systematische Strukturierung zu erhalten.
- Ein raumbezogenes Objekt (Geoobjekt) besteht aus Geometrie- und Sachdaten. Es wird festgelegt, mit welchem Datentyp die Geometrieinformation abgebildet wird, in Raster- oder Vektorform oder kombiniert (hybrid).
- Mit den Geometriedaten werden Position und Form eines Objekts beschrieben. Wesentliches Merkmal der Geometrie ist der Raumbezug, der üblicherweise durch das zugrunde gelegte Koordinatensystem festgelegt wird.
- Grundlegende topologische Beziehungen sind notwendig, um Geodaten strukturiert zu speichern. Je umfassender die Topologie raumbezogener Daten erfasst wird, desto größer sind die Möglichkeiten der Datenanalysen.
- Die Sachdaten umfassen sämtliche Eigenschaften und Beschreibungen von Objekten. Im Zeitalter von Multimedia zählen hierzu auch Textdokumente, Bilder, Video- und Tonsequenzen.
- Im konzeptionellen Schema wird auch festgelegt, in welcher Form die dritte Dimension (Höhe) berücksichtigt wird, ob als ergänzende Angabe in Form eines Attributs oder als digitales Höhenmodell.
- Auch der Umfang der Datenerfassung, die Aktualisierung und Fortführbarkeit der Datenbestände muss bereits in der Konzeptionsphase kalkuliert werden.

verwendete Terminologie	*Englisch*	*Bedeutung*
Geoobjekt	*spatial object*	*Abstraktion einer räumlichen Entität*
Geometrieelement	*saptial feature*	*beschreibt räumliches Merkmal*
Attribut / Eigenschaft	*attribute*	*beschreibt thematisches Merkmal*
semantische Assoziation	*relationship*	*thematische Beziehung*
topologische Assoziation	*spatial association*	*räumliche Verknüpfung (Inzidenz)*
topologische Relation	*spatial relationship*	*räumliche Beziehung (top. Operator)*
Merkmal	*feature*	*besondere Eigenschaft, Beziehung*

5 Logisches Datenmodell (logischer Entwurf)

Ziel des logischen Entwurfs ist es, das semantische Datenmodell mit den Datenstrukturen eines konkreten Herstellersystems (GIS-Software) optimal zu adaptieren. Dabei erfolgt bereits bei der Auswahl einer GIS-Software auch die Festlegung auf eine Modellierungstechnik!

Besonders die Möglichkeiten der geometrisch/topologischen Modellierung sind sehr unterschiedlich ausgeprägt. Die Modellierung statischer topologischer Beziehungen erlauben nur einige wenige GIS-Produkte.

Große Unterschiede bestehen auch in der Datenverwaltung. Bei einfacheren GIS-Produkten wird meist nur das Dateiverwaltungssystem zur persistenten Speicherung der Geodaten verwendet, wobei häufig zusätzlich ein ODBC-Zugriff auf Standarddatenbanken für die Integration von Sachdaten möglich ist. Für große Geodatenbestände ist wegen der Datensicherheit, der Datenintegrität und der parallelen Zugriffe für die Datenänderung die gemeinsame Speicherung von Geometrie- und Sachdaten in einem Datenbanksystem erforderlich. Hierfür werden von einigen GIS-Herstellern sogenannte Extensionen (Erweiterungen) zu Standarddatenbanken wie z.B. Oracle-Spatial (Oracle), SICAD-GDB (AED-SICAD) oder ArcSDE (ESRI) angeboten.

Fachanforderung (reale Welt)
Anforderungsanalyse
Konzeptioneller Entwurf (UML, ERM) Semantisches Datenmodell
Auswahl eines Herstellersystems
Logischer Entwurf (Datenstruktur, DDL) Logisches Datenmodell
Physischer Entwurf (Implementierung)

Die Umsetzung des konzeptionellen in das logische Schema hängt davon ab, welche Modellierungsmöglichkeiten die GIS-Software und das eingesetzte Datenbanksystem unterstützen.

Der logische Entwurf ist **herstellerabhängig** und beschreibt, wie das Geodatenmodell auf Ebene der ausgewählten Produkte realisiert werden kann.

Bereits vor Festlegung auf eine bestimmte GIS-Software sollte deshalb untersucht werden, welche Modellierungsmöglichkeiten zur Verfügung stehen.

Problem: fehlende Interoperabilität

Bei der Konzeption von Informationssystemen lässt sich strukturelles Wissen über die Umwelt in Termen von Gegenständen (Entitäten) mit ihren Eigenschaften und Beziehungen (Assoziationen) ausdrücken. Diese Strukturen kann man allgemein mit objektorientierten semantischen Datenmodellen beschreiben. Für den logischen Entwurf und seine Realisierung mit einer konkreten GIS-Software müssen jedoch die vom Produkt unterstützten Datenstrukturen eingesetzt werden.

Anmerkung: Für die Anwendung ist es in erster Linie nicht ausschlaggebend, nach welchen Konzepten die Entitäten einer Fachwelt im jeweiligen Datenbanksystem intern als Objekte modelliert sind. Vielmehr ist für primär von Bedeutung, dass die Datenbank die für seine Anwendungen erforderlichen Objekte als logische Einheit, performant und mit dem nötigen Detailierungsgrad jederzeit bereitstellen kann und die Objekte persistent und konsistent verwaltet.

Logisches Datenmodell

Als Grundlage für den logischen Entwurf mit einer konkreten GIS-Software sind die Fragen zu klären:

- Welche Datenstrukturen / Operationen stellt die GIS-Software zur Verfügung ?
- Kann der konzeptionelle Entwurf mit dieser GIS-Software optimal adaptiert werden, sodass ein effizientes und performantes Datenmodell entsteht ?

Die Geoobjekte des konzeptionellen Schemas sind mit den konkreten Datenstrukturen, die in der Software verfügbar sind abzubilden und zu implementieren. Auf Ebene der logischen Modelle wird häufig eine andere Terminologie verwendet als auf Ebene der semantischen Modellierung, die weitgehend fachlich orientiert ist:

Geoobjekt – Feature
Geometrieelement – spatial Feature
Attribut – Property

Semantisches Datenmodell	Logisches Schema	Anmerkung
Geoobjekt: **Eindeutig identifizierbare Einheit** des modellierten GeoRaumes aus fachlicher Sicht mit seinen charakteristischen Eigenschaften, entsprechend den Vorgaben der Objektklasse.	Feature: **Besonderes Merkmal** eines Geoobjektes, das seine Form und Lages beschreibt. Wird häufig synonym für Geo-objekt benutzt, obwohl der Begriff Feature dem Objektbegriff inhalt-lich nicht entspricht.	Feature = ***Besonderheit, Merkmal*** Feature type bezeichnet den Geometrietyp einer Klasse (2D: Punkt | Linie | Fläche). Geoobjekte mit gleichem feature type kann man zusammenfassen, man bezeichnet diese Datenor-ganisation dann als feature class.
Geometrieelement: Punkt, Linie, Fläche Grundlegende geometrische Datenstruktur zur Modellierung der **metrischen Merkmale** und **topologischen Beziehungen** (Form, Lage, Nachbarschaft) von Geoobjekten.	Spatial Feature: Grundlegende geometrische Datenstruktur, die in einer GIS-Software zur Verfügung steht. Heute noch sehr unterschiedlich, zukünftig Standardisierung / Normung (→ Spatial Schema)	Spatial Feature = ***räumliches Merkmal*** ***Oberbegriffe für geometrische Datenstrukturen*** Spatial Schema *(ISO 19107)* Simple Feature *(ISO 19125)*
Attribut: Grundlegende Datenstruktur zur Beschreibung der **thematischen Merkmale** (Sachdaten) und **sematischen Beziehungen** von Geoobjekten.	Property: Eigenschaft eines Features, dessen Speicherung im logischen Modell vorgesehen ist.	Attribut = ***charakteristische Eigenschaft*** Proberty = ***Eigenschaft***

Feature Class: Organisationseinheit (Objektthema), die Geoobjekte – Instanzen einer Klasse – mit gleichem Geometrietyp (feature type) enthalten (Objekt ≠ Feature).

Nachfolgend sollen an einigen Beispielen Grundzüge herstellerspezifischer Konzepte für logische Geodatenstrukturen aufgezeigt werden. Es besteht heute noch immer ein Unterschied zwischen den im konzeptionellen Schema formal dargestellten Erfordernissen und den für die Implementierung verfügbaren logischen Datenstrukturen.

Der derzeit stattfindende Standardisierungs- und Normungsprozess hat u.a. das Ziel syntaktisch interoperable Geodatenstrukturen für GIS-Produkte bzw. den Austausch von Geoinformation vorzugeben.

siehe: Open Geospatial Consortium (OGC)

Geometrieelemente (spatial feature)

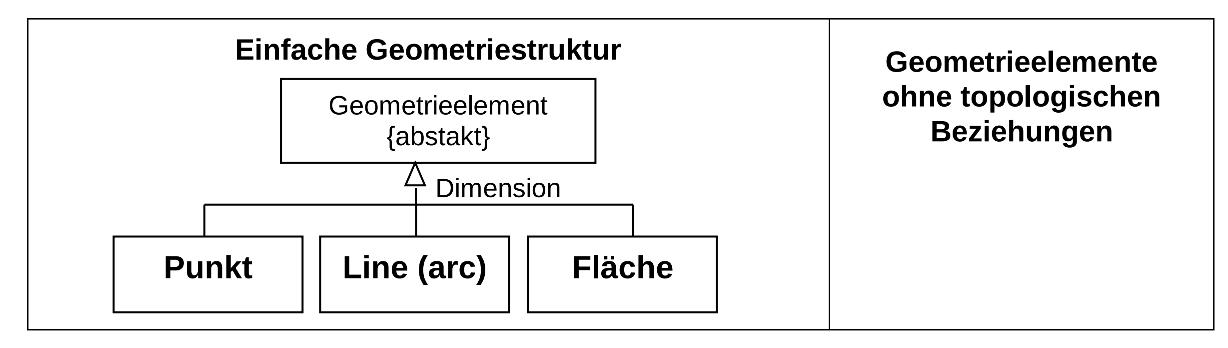
Die Datenstrukturen für die geometrischen Merkmale unterscheiden sich in den GIS-Produkten sehr stark. Viele stellen nur einfache Geometriedatenstrukturen (analog Simple Feature) ohne Topologie zur Verfügung. Andere ermöglichen die gemeinsame Abbildung von Geometrie und Topologie in den Geometrieelementen.

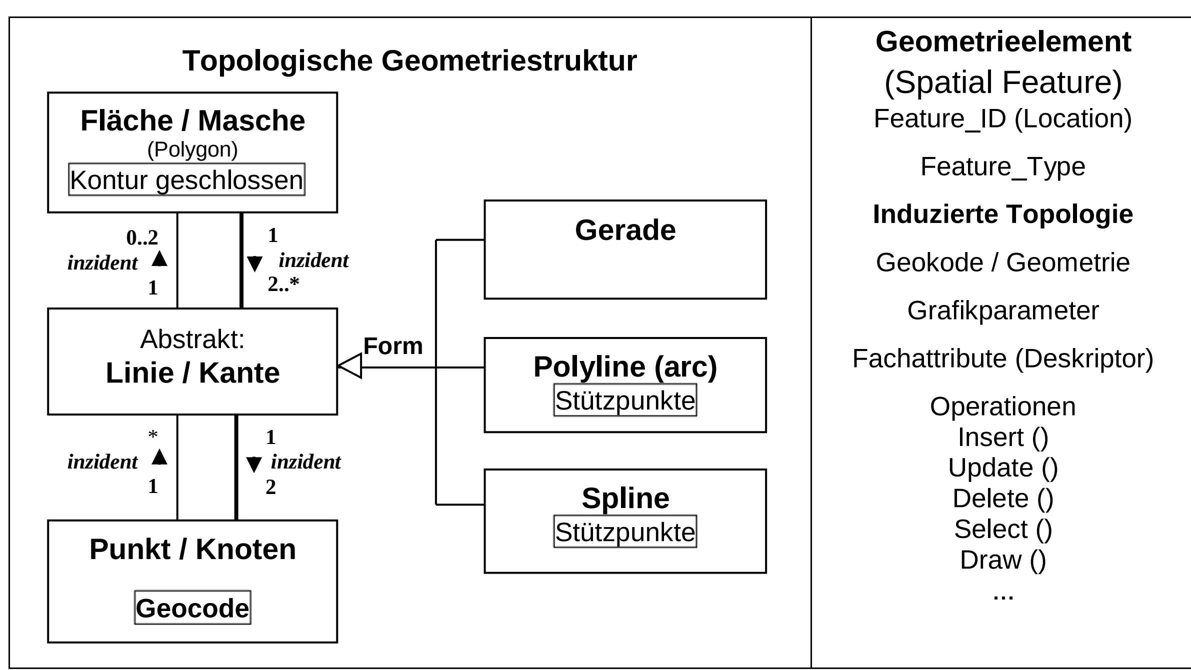
Geometrieelemente (Definitionen ESRI) stehen in folgenden Beziehungen:

Spatial Feature	Topology Element	Relation
Polygon (Area)	Face	Edges define a polygon boundary Nodes are located at edge intersections Vertices define the shape of edges
Line (Arc)	Edge	An edge is defined by a line Two nodes define the endpoints of an edge Vertices define the shape of edges
Point	Node	Points have node behavior when coincident with other features in a topology

Es werden unterschieden:

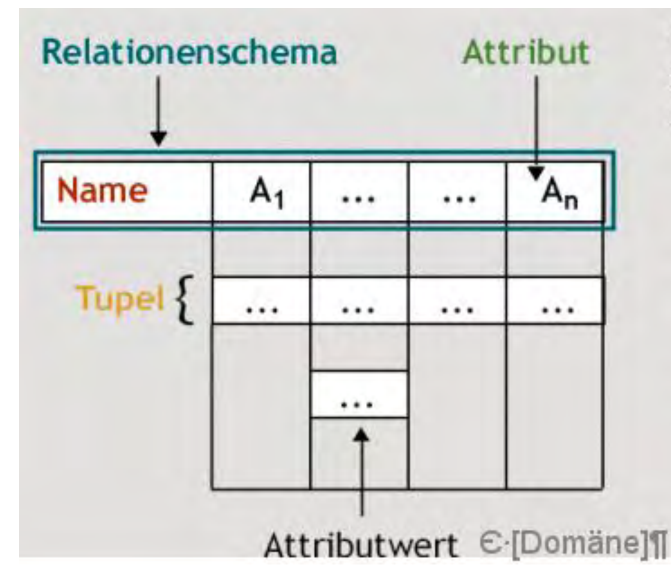
- eigenständige Punktobjekte, im Sinne eines topologischen Knotens (node)
- Stützpunkte (vertices), die nur die Form von Linien approximativ festlegen und als Stützpunktfeld an Linienelementen gespeichert werden.





Attribute

Die Thematik umfasst alle fachlichen Merkmale einer Klasse. Sie kann durch die Werte einer Menge von Attributen beschrieben werden, die die Objekte charakterisieren. Attribute können nominale, ordinale, kardinale oder metrische Werte enthalten. Die Organisation und Verwaltung der Attribute erfolgt meist in Tabellen, entsprechend dem relationalen Datenbankmodell. In einer Tabelle wird die Bedeutung eines Attributs durch den Namen der Spalte festgelegt, die konkrete Ausprägung einer Struktur, das Tupel (Instanz einer Tabelle), ergibt sich aus den Werten einer Zeile.



Eine relationale Datenbank besteht aus einer oder mehrerer **Tabellen** (**Relationen**)

- Die Spalten einer Tabelle heißen Attribute und sind typisiert
- Die Zeilen einer Tabelle heißen Tupel
- Eine Relation erhält einen Relationennamen
- Vorgabe der Attribute legt die Struktur einer Tabelle fest: Relationenschema
- Beziehungen zwischen den Tabellen werden als Relationship bezeichnet.

Nach der von *Bill/Fritsch* angegeben **Dreiebenen-Hierarchie** für die vollständige Konfiguration von Geoobjekten bildet die **attributive Beschreibung eines Geoobjekts die oberste Ebene.** Bei Geoobjekten kann man deshalb ein Tupel als „Objektrumpf" verwenden, der die semantische Struktur eines Geoobjekts darstellt und auch die OID enthält. Die Eindeutigkeit der Tupel kann mittels Primärschlüssel gewährleistet werden. Die Primärschlüsselwerte sind mit geeigneten Generierungsvorschriften so abzuleiten, dass sie den inhaltlichen Anforderungen eines OID entsprechen.

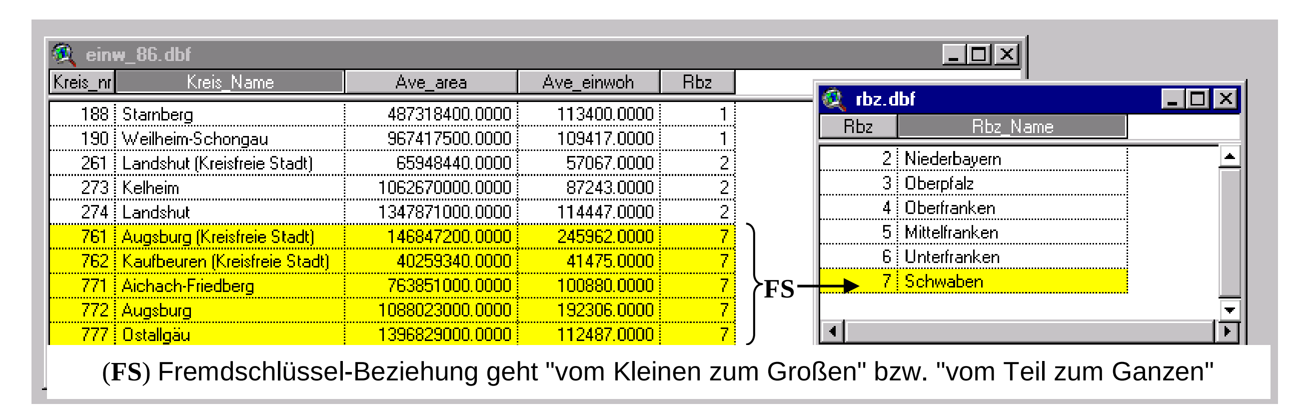
Primärschlüssel werden von einem oder mehreren Attribut(en) gebildet, sie werden zur eindeutigen Identifizierung beim Zugriff verwendet, sie realisieren die Entity Integrität. Primärschlüssel-Attribute müssen Werte haben („not Null"), der Primärschlüssel muss die Bedingung „unique" erfüllen.

Zur Generierung einer OID werden häufig sogenannte Surrogate verwendet, dies sind systemweit eindeutige, persistente Bezeichner, die unabhängig vom physischen Speicherort sind. Sie können z.B. aus Aggregaten von Datums- und Zeitangaben der Systemuhr gebildet werden. Die OID ergibt nur eine systemweite, **anonyme Identität** eines Geoobjekts, sie entspricht seiner „Personalausweisnummer", über die man es jedoch nur schwerlich ansprechen kann. Für die eindeutige logische Ansprache eines Geoobjektes sind deshalb noch zusätzliche Attribute für seine **fachliche Identität** erforderlich, seinen „Personalien" (wer). Der Geokode (Raumbezug) ist die zugehörige Lokalisation, seine „Adresse" (wo), über die man es räumlich selektieren kann.

Die Kapselung der Objektzustände (= Werte der Attribute) wird in GIS meist nicht realisiert, da sie für Datenbankobjekte ohnehin nur bedingt sinnvoll ist. Datenobjekte sollen in vielen Fällen gerade mittels Prädikaten über Teile ihrer Werte in einer Menge von Objekten aufgesucht werden (z.B. ... *where Fläche > 10 and Nutzung = 'AD'*).

Semantische Beziehungen

Die semantischen Beziehungen (Assoziationen) können über Fremdschlüsselbeziehungen realisiert werden, die auf der Referenz von semantisch übereinstimmenden Attributkombinationen in zwei Tupeln beruhen. Sie werden für die Beschreibung von Assoziationen mit Instanzen anderer Klassen verwendet, insbesondere zu Aggregaten oder zu Objekten, die primär keinen Raumbezug haben (z.B. Flurstück → Eigentümer).



einw_86.dbf

Kreis_nr	Kreis_Name	Ave_area	Ave_einwoh	Rbz
188	Starnberg	487318400.0000	113400.0000	1
190	Weilheim-Schongau	967417500.0000	109417.0000	1
261	Landshut (Kreisfreie Stadt)	65948440.0000	57067.0000	2
273	Kelheim	1062670000.0000	87243.0000	2
274	Landshut	1347871000.0000	114447.0000	2
761	Augsburg (Kreisfreie Stadt)	146847200.0000	245962.0000	7
762	Kaufbeuren (Kreisfreie Stadt)	40259340.0000	41475.0000	7
771	Aichach-Friedberg	763851000.0000	100880.0000	7
772	Augsburg	1088023000.0000	192306.0000	7
777	Ostallgäu	1396829000.0000	112487.0000	7

rbz.dbf

Rbz	Rbz_Name
2	Niederbayern
3	Oberpfalz
4	Oberfranken
5	Mittelfranken
6	Unterfranken
7	Schwaben

(**FS**) Fremdschlüssel-Beziehung geht "vom Kleinen zum Großen" bzw. "vom Teil zum Ganzen"

In den meisten GIS werden für die Abbildung von semantischen Assoziationen Fremdschlüsselbeziehungen verwendet.

Fremdschlüssel sind Attribute in einer Tabelle, deren Werte der gleichen Domäne eines Primärschlüssels einer referenzierten Tabelle angehören. Der Fremdschlüssel realisiert eine Inklusionsabhängigkeit, die als referentielle Integrität bezeichnet wird.

Der Fremdschlüssel stellt eine (1:n)-Beziehung zweier Klassen her. Fremdschlüssel bei existenzabhängigen Teile-Beziehungen (Komposition) sind „on delete cascade“ zu deklarieren, damit bei einer Löschung des Ganzen auch die referenzierten Teile gelöscht werden.

Datenstrukturen im Shape-Format (ESRI)

Die bisher dargestellte Art der logischen Modellbildung für Geoinformation, d.h. die Verwendung verschiedener Datenstrukturen für Geometrie und Sachdaten, entspricht den in vielen GIS-Produkten verfügbaren Möglichkeiten zur die Bildung von Geoobjekten. Für die Zusammenfassung der verschiedenen Datenstrukturen zu einem Geoobjekt wird meist zusätzlich noch eine spezielle systemunterstützte Datenstruktur „Geolink“ verwendet (z.B.: Shape-Format: shx, SICAD: location).

Als Geolink (Spatial Link) wird eine spezielle 1:n-Relation <OID,{FID}> bezeichnet, mit der die beiden Datenstrukturen Geometrieelemente (spatial Feature) und Attribute (Tupel) zur Bildung von Geoobjekten, als Informationseinheiten, in Beziehung gesetzt werden. In Analogie zu den topologischen Beziehungen kann man den Geolink auch als eine relationale Inzidenzbeziehung zwischen den Objektteilen Tupel (Attribute) und Geometrieelemente betrachten, die mittels Geolinks realisiert wird.

Die Verwaltungseinheit für die Geometriedaten ist das Thema. Die Datenverwaltung der Themen erfolgt in Dateien. Ein Thema im Shape-Format (Objektthema) besteht jeweils aus drei Dateien, der Shape-Datei mit der Geometrie, der Attribut-Datei mit den "Attributen des Themas" im dBASE-Format und der Verbindungsdatei (Geolink), die die (1:1)-Beziehungen zwischen Geometrie- und Sachdaten herstellt. Ein Shape-Thema wird auch häufig auch als Feature Class bezeichnet, da es insbesondere auch den Geometrietyp der enthaltenen Objekte festlegt (Punkt|Linie|Fläche).

Zusätzliche Sachdaten können als dBASE-Dateien oder in Datenbanken (z.B. Access) verwaltet werden. Für die Tabellen mit Sachdaten im dBASE-Format können keine Integritätsbedingungen und keine "ist Teil von" - Beziehungen definiert werden.

Elementare Datenstrukturen (ArcView GIS, Shape-Format)

- Geometrieelemente (Punkt | Linie | Fläche), ohne Topologie („simple feature“)
- Attribut-Tupel (Sachsatz)
- Rastermatrix (Bild, Grid)

Datenorganisation

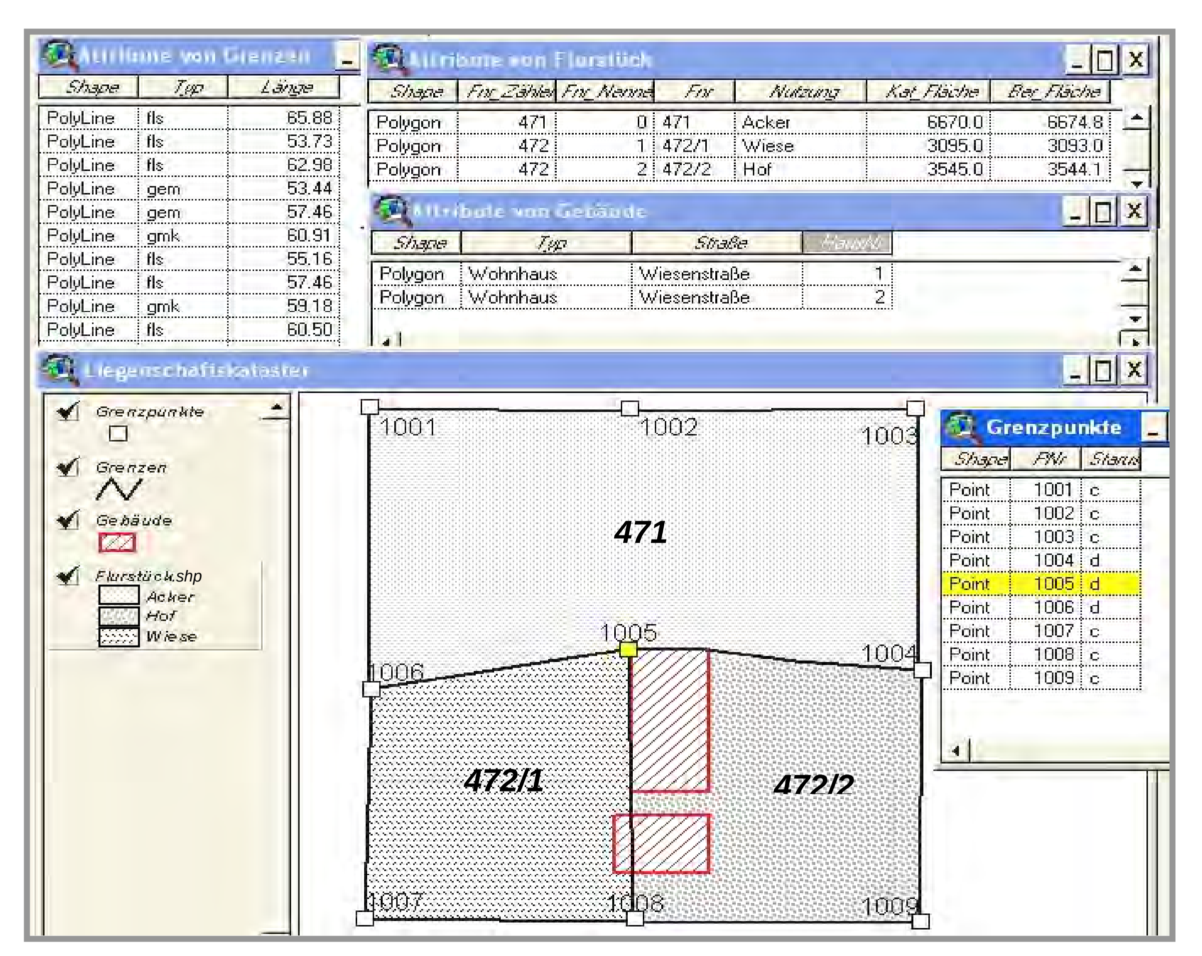
- Thema ::= <Objektthema = Feature Class | Bildthema>
  Ein Thema kann nur einen Typ von Geometrieelementen enthalten und ist mit anderen Themen nur über den Raumbezug verknüpft.
- Objektthema ::= <Name,{Geoobjekt(Typ ::= Punkt | Linie | Fläche)},[join_Tabelle]>
- Bildthema ::= <Name,Rastermatrix> └ {Geoojekt(Typ_Punkt) | Geoojekt(Typ_Linie) | ..}
- Tabellen ::= <Name,{Attribut-Tupel}>

Objektstruktur

- Geoobjekt ::= <FID,Klasse-Thema,Geometrieelement,Geolink,Tupel>
  Datenstruktur für Geoobjekte = (typ_shp<>shx<>dbf)
  Die Eigenschaften der Klasse ergeben sich aus dem Typ der Geometrie und den Attributen des Themas (FID=OID)

Verbindungen (Link)

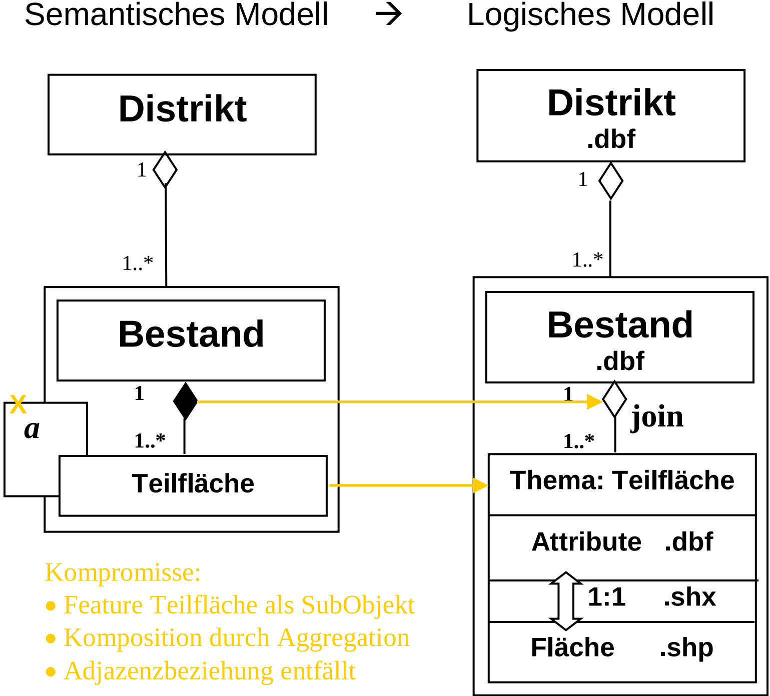
- Geolink (shx) ::= <FID,SID>, 1:1 Beziehung zwischen Geometrieelement/Tupel
- Join, Relate ::= <Tabelle_Thema – join – Tabelle> Verbindung, Verknüpfung
- Hotlink ::= <Tabelle_Thema(Attribut)> Verweis auf allgemeine Dokumente

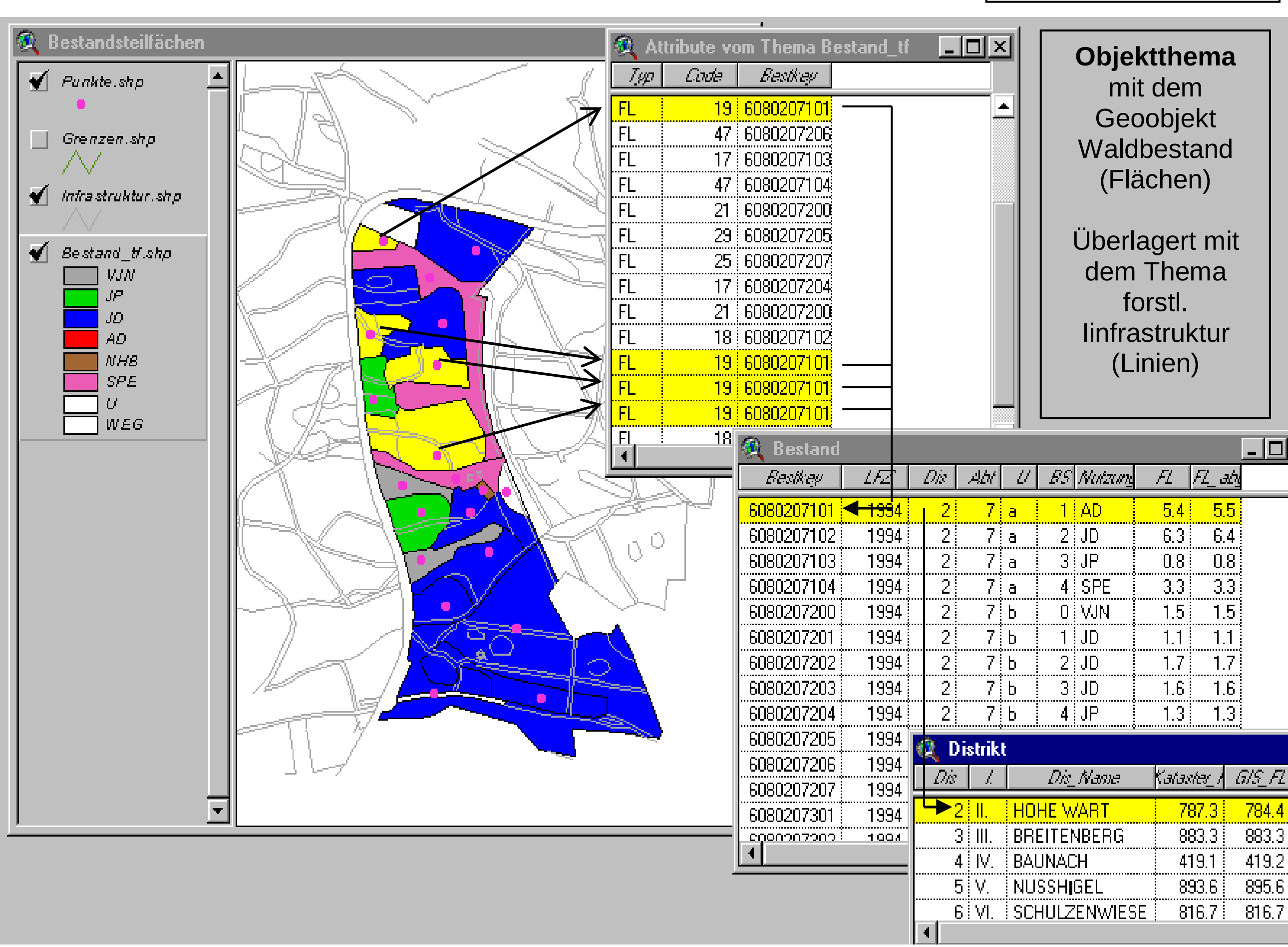


Shape-Format: Beispiel Geoobjekt Waldbestand

Umsetzung des konzeptionellen Schemas für das Geoobjekt Bestand in ein logisches Modell im Shape-Format.

Anmerkung: Aus fachlicher Sicht ist der Bestand das kleinste Objekt im Themenbereich Forstbetrieb. Bestandsteilflächen, die nötig sind, um das logische Schema im Shape-Format zu realisieren, haben fachlich keine Bedeutung. Sie müssen jedoch als Subobjekte eingeführt werden, da die direkte Bildung des Geoobjekts Bestand nicht möglich ist. Die Verbindung der Subobjekte mit dem Bestand erfolgt mittels eines Join (Verbindung).





***Shape-Format:** Das Dateiformat **Shape** ist ein von ESRI ursprünglich für ArcView GIS entwickeltes simple feature Format für Geodaten, das in mehreren Dateien gespeichert wird. Es hat sich mittlerweile zu einem Industriestandard im GIS-Umfeld entwickelt, da es ein recht einfaches, aber für viele Zwecke ausreichendes Format für objektstrukturierte Vektordaten darstellt. Es wird praktisch von allen GIS-Produkten und Datenprovidern unterstützt.*

Topologischer Beziehungen (Inzidenzen)

Für die Kodierung von statischen topologischen Beziehungen im Datenmodell gibt es zwei Möglichkeiten (siehe ISO 19107, Spatial Schema):

- **Separat** in eigenen Inzidenzmatrizen. Aus der Geometrie wird mit speziellen Funktionen die Topologie abgeleitet und in Inzidenzmatrizen verwaltet.

  Konsequenz: *Die topologische Information ist projektbezogen nur für einen Teil der Datenbasis (Coverage, Thema) verfügbar und muss bei Änderungen überarbeitet werden.*
- **Integriert** im Geometrieelement. Die Inzidenzen werden im Geometrieelement gespeichert, das die Topologie induziert. Die Speicherung erfolgt in Form von Verweisen (spatial links, location) auf die inzidenten Geometrieelemente.

  Konsequenz: *Die Geodatenbasis kann beliebig groß werden (offen), da die Topologie gemeinsam mit der Geometrie objektbezogen verwaltet wird. Bei Änderungen müssen nur die betroffenen Geometrieelemente angepasst werden.*

Die Darstellung der Inzidenzen kann in einem Beziehungsgraphen erfolgen, der i.d.R. ein Netzwerk bildet, da zwischen den Geometrieelementen m:n-Relationen bestehen; von einem Knoten können mehrere Kanten abgehen und eine Kante hat einen Anfangs- und Endknoten (siehe oben).

Topologische Beziehungen dienen der effizienten Verarbeitung von Geoobjekten, sie entsprechen den äußeren Beziehungen von Geoobjekten auf Ebene gemeinsamer Knoten und Kanten, den Inzidenzen und Adjazenzen. Man braucht sie zur:

- **Vermeidung redundanter Geometrieelemente**
- Nachbarschaftsbeziehungen über gemeinsame Knoten und Kanten (Kollektive)
- Netzverfolgung, Routing mit kombinatorischen Algorithmen
- Organisation von Geotransaktionen
- Prüfen von Konsistenzbedingungen für Inzidenzen

Separate Speicherung der Topologie (Beispiel ARC/INFO coverage)

Die Beziehungen zwischen den topologischen Elementen werden in separaten Inzidenzmatrizen verwaltet, wobei auch die Orientierung der Kanten vermerkt wird. Die Inzidenzmatrizen beschreiben den gerichteten planaren Graphen der Geometrie.

Die **Topologie** wird separat in Inzidenzmatrizen gespeichert:

- ***Arc-Node-List*** = Kanten-Knoten-Inzidenz → zwei inzidente Knoten zu einer Kante, durch die **from/to**-Angabe ergibt sich die Kanten-Knoten-Richtung
- ***Left-Right-List*** = Kanten-Maschen-Inzidenz → zwei inzidente Maschen einer Kante, durch die **left/right**-Angabe ergibt sich Kanten-Maschen-Richtung
- ***Polygon-Arc-Lis***t = Manschen-Kanten/Knoten-Inzidenz mit den Kanten und Knoten einer Polygonkontur (polygon boundary oder border polygon)

Die **Geometrie** (Metrik) ist in einer weiteren Liste, der ***Arc-Coordinate-List*** gespeichert, die über die ***FID*** der Arcs referenziert wird und ein Koordinatenfeld enthält, das mit der Koordinate des ***A-Node*** beginnt, n-formbestimmende (n = 0, 1, ...) ***Vertices*** beinhaltet und mit der Koordinate des ***E-Node*** abschließt.

Punkte (Koordinaten) aus topologischer Sicht unterscheiden !
Node = Punkt, der Anfangs-/Endpunkt einer Kante ist (Knoten)
Vertex = Stützpunkt, der nur die Form einer Kante beschreibt

Geometrie

Arc-Coordinate-List

Border Polygon

Polygon-Arc-List

Topologie

Masche (Polygon) — Right-Poly / ***Left-Right-List*** / Left-Poly — ***Kante (Arc)*** — To-Node / ***Arc-Node-List*** / From-Node — ***Knoten (Node)***

Thematik

Polygon Attribute | ***Arc Attribute*** | ***Node Attribut***

(0,0) x y

Kanten: 1, 2, 3, 4, 5, 6 — Maschen: (1), (2), (3), (A) — Knoten: △1, △2, △3, △4

Beispiel: ArcInfo-Datenstruktur

Arc	Coordinates (x,y)
1	**3.6,0.0** 6.5,0.0 6.5,5.0 **4.6,5.0**
2	**4.6,5.0** **0.7,5.0**
3	**0.7,5.0** 0.0,5.0 0.0,0.0 **3.6,0.0**
4	**3.6,0.0** 3.1,0.6 2.8,1.2 2.6,1.6 2.7,2.2 2.9,2.9 **3.2,3.6**
5	**3.2,3.6** 3.5,3.8 3.8,4.5 **4.6,5.0**
6	**3.2,3.6** 2.8,3.8 1.6,4.2 **0.7,5.0**

Arc-Coordinate-List
Koordinaten der Stützpunkte (vertices) und Knoten

Polygon	***Arcs***	***Nodes***
⌐A	***+1,+2,+3***	△1 △2 △3
(1)	***+3,+4,+6***	△1 △4 △3
(2)	***+1,-5,-4***	△1 △2 △4
(3)	***+2,-6,+5***	△2 △3 △4

Polygon-Arc-List
Maschen-Kanten-Knoten-Inzidenz

Adjazente Maschen

Arc	***Left-Polygon***	***Right-Polygon***
1	(A)	(2)
2	(A)	(3)
3	(A)	(1)
4	(2)	(1)
5	(2)	(3)
6	(3)	(1)

Left-Right-List
Kanten-Maschen-Inzidenz

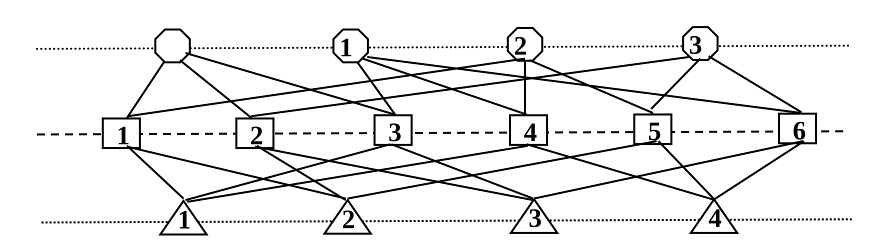
Adjazente Knoten

Arc	***From-Node***	***To-Node***
1	△1	△2
2	△2	△3
3	△3	△1
4	△1	△4
5	△4	△2
6	△4	△3

Arc-Node-List
Kanten-Knoten-Inzidenz

innere Struktur („selbstbezogen“) äußere Struktur – topologische Beziehungen („fremdbezogen“)

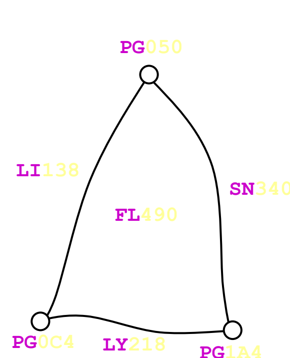
Beziehungsgraph zur Geometrie, Abbildung der 3 Listen (ohne Orientierung)



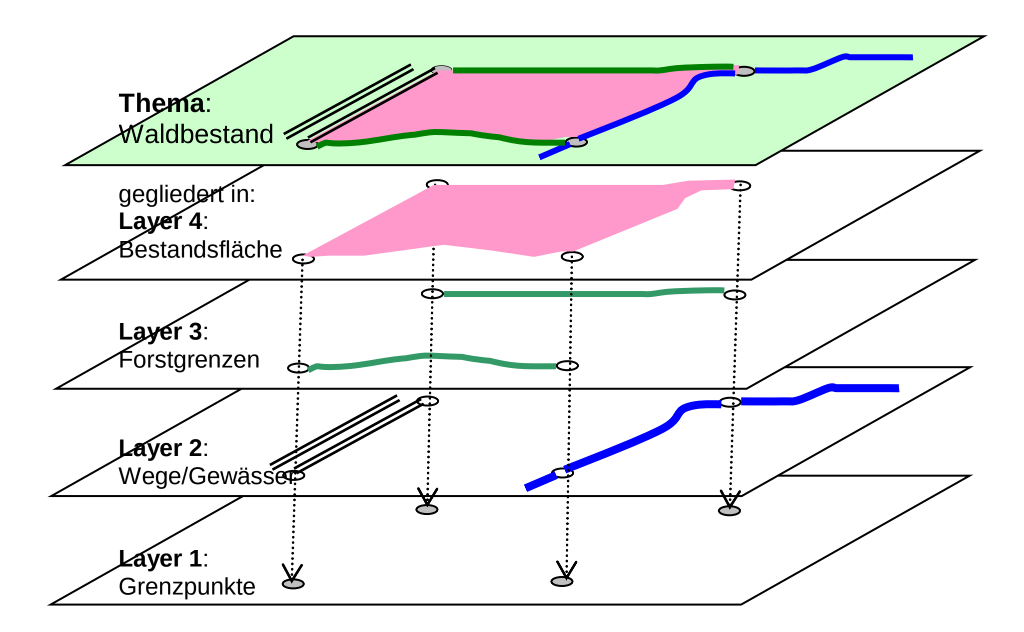
Der Graph der Inzidenz- und Adjazenzbeziehungen zwischen Knoten, Kanten und Maschen ergibt ein Netzwerk (m:n-Relation).

Integrierte Speicherung der Topologie (Beispiel SICAD-GDB)

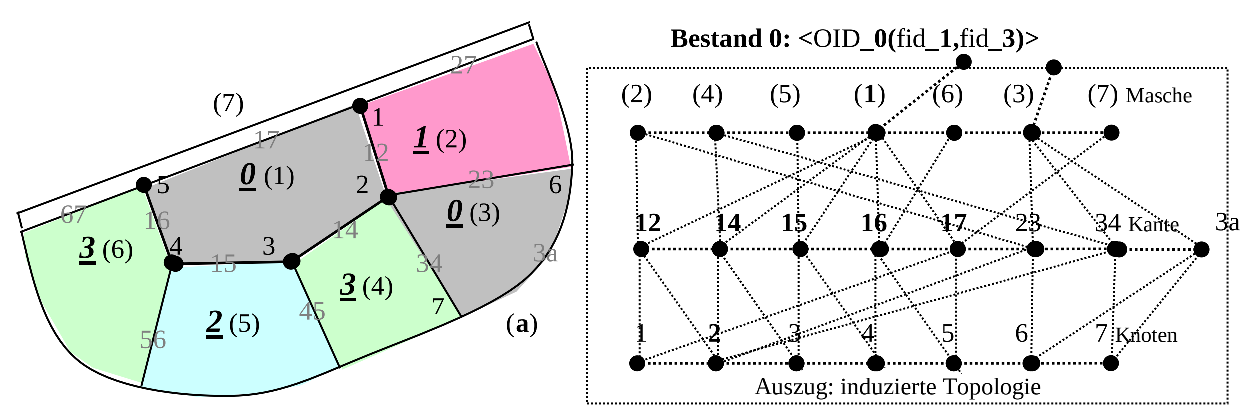
Für die Abbildung der Geometrie stehen die Geometrieelemente (saptial Feature) Punkt (PG), Linie (LI, LY, SN, ...) und Fläche (FL) zur Verfügung. Jedes Geometrieelement erhält eine datenbankweite Identität (Location), die hier als Feature-Identifier (FID) bezeichnet wird. Die Inzidenzbeziehungen werden im Geometrieelement, das die topologische Beziehung indiziert, lokal mittels FID gespeichert.



ELEMENTTYP PG (Punkt PG050) EBENE 1 X 4362225.88 Y 5563298.70 PRIORITAET 0.000000 PKZ G0 PNR 1001	ELEMENTTYP LI (Gerade LI138) EBENE 2 PRIORITAET 1.000000 SM 1 ST 1 TXT 'StFoV'
Topologie / Inzidenz-Beziehungen PG 050 LI 138 SN 340 PG 0C4 LI 138 LY 218 LI 138 StFoV FL 490 PG 1A4 LY 218 SN 340 LY 218 Distrikt FL 490 SN 340 Abteilung FL 490 FL 490 VJN (vereinfachte FID)	ELEMENTTYP LY (Polyline LY218) EBENE 3 FLD .4362979 E 7 .5563499 E 7 ... .4363177 E 7 .5563094 E 7 PRIORITAET 1.000000 ST 2 SM 2 TXT 'Distrikt'
ELEMENTTYP FL (Fläche FL490) EBENE 8 X .4362784 E 7 Y .5563310 E 7 PRIORITAET -7.000000 W 0.000000 SA 2.500000 FCO 21 NAM VJN FLA 356775.476 Geoelement::= < FID, Typ, Layer, {Inzidenz}, [Geokode, {Formparameter}], [{Grafikparameter}], [{Deskriptorattribut}] >	ELEMENTTYP SN (Spline SN340) EBENE 4 ZSP FREI LEN 1104.277448 WAN -82.385826 WEN -162.782787 FLD .4362254 E 7 .5563215 E 7 ... .4362914 E 7 .5562965 E 7 PRIORITAET 1.000000 ST 3 SM 3 TXT 'Abteilung'



Erfassung und Speicherung der Inzidenzbeziehungen



Liniennummer ergibt sich aus den Nummern adjazenter Teilflächen

Damit bei der Datenerfassung die Geometrieelemente topologisch richtig miteinander verkettet werden, müssen die eingesetzten Verfahren aus dem Sachverhalt der Anwendung heraus (Maßstab, Datendichte, Verfahrensablauf) optimal parametrisiert sein.
Besonders wichtig sind die Einstellungen der Optionen für die Punkt- /Linienidentifikation und das „Auftrennverhalten“ bei der Erzeugung neuer Punkt oder Linienelemente.

Die Steuerung der Verkettungsoperationen kann durch folgende Optionen erfolgen:

- Toleranz zur Identifikation von Punkten und Linien (Elementidentität)
- Toleranz zur Identifikation von Linienstützpunkten (Vertex)
- Auftrennverhalten von vorhandenen Linien innerhalb der Toleranz beim Erzeugen neuer Punkte oder Auftrennung von neuen Linien an Punkten, die innerhalb der Toleranz vorhandenen sind
- Linientypen, auf die sich das eingestellte Auftrennverhalten bezieht

FID	ELEMENT		Inzidenzen
49C	FL	**(1)**	FL 49C
50	PG	**1**	LIPG 138 LIPG 434 LIPG 590
C4	PG	**2**	LIPG 138 LIPG 214 LIPG 698 LIPG 7A0
138	LI	**12**	LI 138 FLLI 49C FLLI F1C
1A0	PG	**3**	LIPG 214 LIPG 2F0 LIPG 8AC
214	LI	**14**	LI 214 FLLI 49C FLLI 101C
27C	PG	**4**	LIPG 2F0 LIPG 3CC LIPG 988
2F0	LI	**15**	LI 2F0 FLLI 49C FLLI 109C
358	PG	**5**	LIPG 3CC LIPG 434 LIPG A90
3CC	LI	**16**	LI 3CC FLLI 49C FLLI 111C
434	LI	**17**	LI 434 FLLI 49C FLLI 119C
51C	PG		LIPG 590 LIPG B24 LIPG C0C
590	LI	**27**	LI 590 FLLI F1C FLLI 119C
624	PG		LIPG 698 LIPG C0C LIPG CA0
698	LI	**23**	LI 698 FLLI F1C FLLI F9C
72C	PG		LIPG 7A0 LIPG CA0 LIPG D34
7A0	LI	**34**	LI 7A0 FLLI F9C FLLI 101C
838	PG		LIPG 8AC LIPG D34 LIPG DC8
8AC	LI	**45**	LI 8AC FLLI 101C FLLI 109C
914	PG		LIPG 988 LIPG DC8 LIPG E5C
988	LI	**56**	LI 988 FLLI 109C FLLI 111C
A1C	PG		LIPG A90 LIPG B24 LIPG E5C
A90	LI	**67**	LI A90 FLLI 111C FLLI 119C
B24	LI	**7a**	LI B24 FLLI 119C
C0C	LI	**2a**	LI C0C FLLI F1C
CA0	LI	**3a**	LI CA0 FLLI F9C
D34	LI	**4a**	LI D34 FLLI 101C
DC8	LI	**5a**	LI DC8 FLLI 109C
E5C	LI	**6a**	LI E5C FLLI 111C
F1C	FL	**(2)**	FL F1C
F9C	FL	**(3)**	FL F9C
101C	FL	**(4)**	FL 101C
109C	FL	**(5)**	FL 109C

I↑ dargestellt sind Knoten-Kanten-Inzidenz

Bei optimaler Einstellung der Optionen (Parameter) für die Identifikation und Verkettung wird die topologische Struktur praktisch ohne Mehraufwand für die Digitalisierer bei der Erfassung der Geometrieelemente systemunterstützt aufgebaut, d.h. die wichtige topologische Grundstruktur zweidimensionaler Flächengeometrien wird durch das System „weitgehend automatisch“ in den Geometrieelementen induziert. Abbildung der Inzidenzen erfolgt in den Geometrieelementen mittels Feature_ID der referenzierten Elemente.

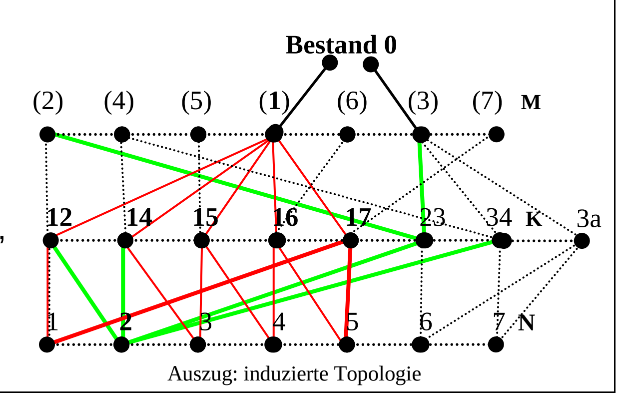
Die toplogischen Beziehungen werden im jeweiligen Geometrieelement gespeichert, wodurch sie als Merkmale von Geoobjekten objektstrukturiert verwaltet werden.

Die topologischen Beziehungen werden lokal in den Geometrieelementen mittels Strukturketten abgespeichert, durch die sich die Inzidenzbeziehungen abbilden lassen. Die Strukturketten enthalten Verweise zwischen Oberelementen (Master) und Unterelementen (Detail). Sie gehen immer von einem Master aus, führen über die zugehörigen Details und enden wieder am Master (Ringstruktur). Mit dieser Datenstruktur können folgende Master/Detail-Beziehungen zwischen den Geometrieelementen abgebildet werden, denen die folgenden topologischen Beziehungen entsprechen:

Geometrie-elementtyp	Induzierte Topologie	Topologische Strukturen von **Geo-Elementen**	Topologische Beziehungen in **ebenen Flächengraphen**
Fläche	**Masche**	kann Master von (2...n) Linien sein	**Maschen-Kanten-Inzidenz** $K\{12,14,15,16,17\} = \mathbf{I}\downarrow M\{1\}$ (alle Kanten zu adjazenten Maschen)
Linie	**Kante**	kann allgemein Detail von (1|2|n) Flächen sein, in ebenen Flächengraphen genau von 2 benachbarten Flächen oder 1 Fläche und dem Außenraum	**Kanten-Maschen-Inzidenz**, Umkehrung Maschen-Adjazenz $m_2 \ \mathbf{V} \ m_3 \Rightarrow M\{2,3\} = \mathbf{I}\uparrow K\{23\}$
		ist Master von genau 2 Punkten	**Kanten-Knoten-Inzidenz**, Umkehrung Knoten-Adjazenz $n_1 \ \mathbf{\Lambda} \ n_5 \Rightarrow N\{1,5\} = \mathbf{I}\downarrow K\{17\}$
Punkt	**Knoten**	kann Detail von (1|2...n) Linien sein, in ebenen Flächengraphen von mindestens 2 Linien	**Knoten-Kanten-Inzidenz** $\mathbf{K}\{12,14,23,34\} = \mathbf{I}\uparrow N\{2\}$ (alle Kanten mit adjazenten Knoten)

Das Beziehungsmuster der Geometrieelemente hat eine Netzwerkstruktur, die auf den vier topologischen Grundbeziehungen beruht.

Die Strukturketten in den Geo-Elementen adaptieren dieses Beziehungsmuster optimal, die Strukturketten enthalten die benötigten Inzidenz-Beziehungen.



Auszug: induzierte Topologie

Mit diesen Inzidenzen lassen sich sowohl der innere Aufbau als auch die äußeren Beziehungen der Geoobjekte auf Geometrieebene topologisch darstellen, sie ergänzen die metrische Information der Koordinaten um die topologische Information.

Aus der Abbildung ist zu ersehen, dass

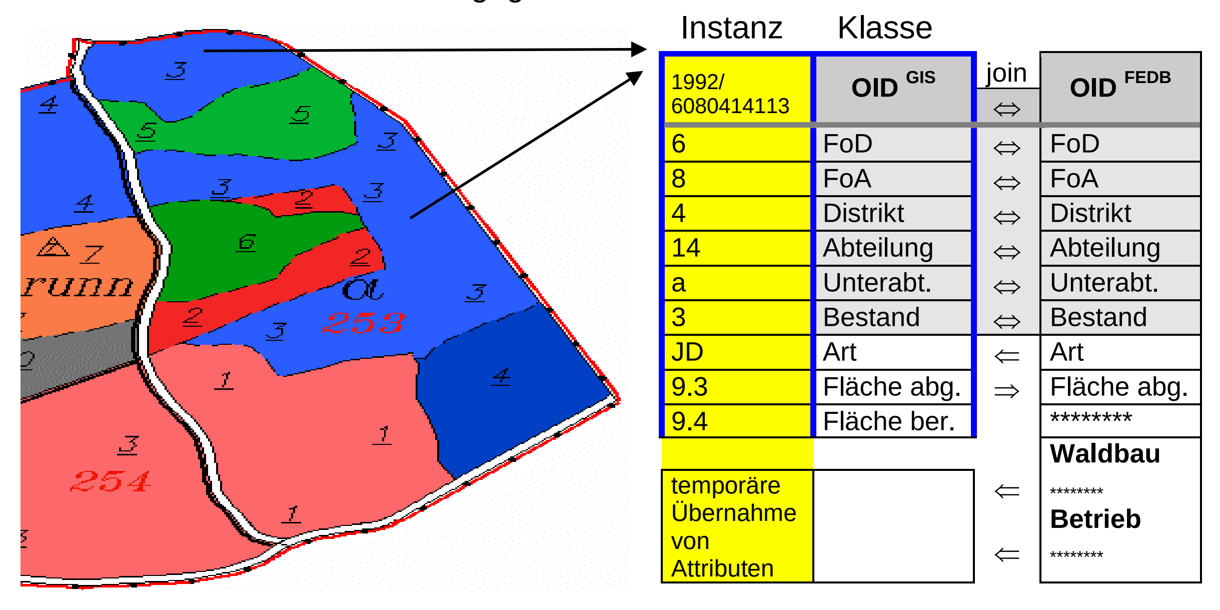
- Inzidenzen zwischen Master und Detail ($\mathbf{I}\downarrow$) die endogene Struktur und
- Inzidenzen zwischen Detail und Master ($\mathbf{I}\uparrow$) die exogene Struktur wiedergeben.

In SICAD stehen spezielle Operatoren <op> ::= ($\mathbf{I}\uparrow$, $\mathbf{I}\downarrow$) für die Auswertung der Inzidenzbeziehungen zur Verfügung, die eine Navigation im Topologiemodell ermöglichen.

Die Ergebnismenge der zu {A} inzidenten Elemente {B} ergibt sich dabei durch Anwendung eines Inzidenzoperators {B} = <op> {A} auf die Menge der Ausgangselemente {A}.

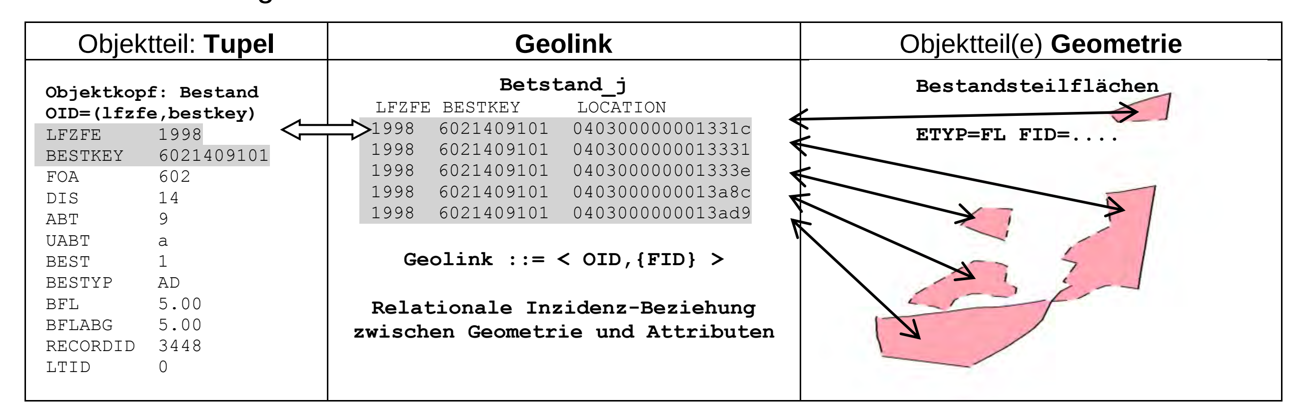
GDB-Format: Beispiel Geoobjekt Waldbestand

Geoobjekte werden aus den Teilobjekten Attribute und Geometrie gebildet, die durch eine Relation Geolink miteinander verbunden werden. Für die Abbildung der Geometrie stehen in der SICAD-GDB Geometrieelemente zur Verfügung, in denen die topologischen Beziehungen mittels Inzidenzen abgebildet werden. Die Realisierung des Teilobjekts Attribute für die thematische Modellierung, erfolgt mit Tabellen. Für Geoobjekte werden Tabellen kreiert, deren Tupel den „Rumpf“ (Header) eines Geoobjekts darstellt (oberste Modellierungsebene). Der Objektrumpf enthält die OID und Attribute mit der Organisation sowie Attribute mit primär geometrischem Inhalt, dies sind z.B. die berechnete und abgeglichene Bestandsfläche.



Instanz	Klasse: OID GIS	join	OID FEDB
1992/ 6080414113	OID GIS	⇔	OID FEDB
6	FoD	⇔	FoD
8	FoA	⇔	FoA
4	Distrikt	⇔	Distrikt
14	Abteilung	⇔	Abteilung
a	Unterabt.	⇔	Unterabt.
3	Bestand	⇔	Bestand
JD	Art	⇐	Art
9.3	Fläche abg.	⇒	Fläche abg.
9.4	Fläche ber.		********
			Waldbau
temporäre Übernahme von Attributen		⇐	********
			Betrieb
		⇐	********

Für die Konfiguration der Geoobjekte ist es nötig Geometrie und Attribute zu verbinden. Wegen der engen Beziehung zwischen den Objektteilen wird dazu in der SICAD-GDB eine spezielle Relation verwendet, die systemunterstützt erzeugt und gepflegt wird; diese Verbindungsinformation wird als Geolink bezeichnet.



Das erstellte Detailkonzept beschreibt eine abstrakte Klasse „Forstfläche“, aus der durch Spezialisierung die Klassen für die elementaren Geoobjekte Bestand oder Standort abgeleitet werden können. Sie erlaubt - ohne Einschränkung - eine übersichtliche und objektstrukturierte Darstellung der Geoinformation mit allen relevanten Merkmalstypen der Geoobjekte.

SICAD-GDB: *Die GeoDatenbasis von SICAD (AED-SICAD) ist eine speziell für Geodaten entwickelte Datenbankextension, die eine topologische Datenstruktur mit elementbezogenen Inzidenzen zur Verfügung stellt. Die Daten können im Format SICAD-SQD exportiert und importiert werden.*

Operationen, Methoden *(siehe Kapitel 7)*

Die Geodaten bilden das statische Modell (Merkmale), die Operationen das dynamische Modell (Verhaltensmerkmale). Operationen können neue Informationen erschließen, die sich aus der gemeinsamen Analyse von unabhängigen Geodatenquellen ergeben.

Operationen beschreiben die Verhaltensmerkmale von Objekten. In objektorientierten Systemen sind die Operationen als Methoden in den Klassen implementiert, damit bilden Objekte abgeschlossene Pakete aus Daten und den dazugehörenden, problem-orientierten Methoden (Kapselung). Man unterscheidet:

- Konstruktoren / Destruktoren zum Erzeugen und Löschen von Objekten
- Mutatoren zum Ändern von Objekten
- Akzessoren für den lesenden und schreibenden Zugriff auf Attributwerte
- **Funktionen** (Tätigkeiten)

In GIS-Softwareprodukten sind die unterschiedlichen Operationen meist verschiedenen Softwarekomponenten zugeordnet. Es gibt, ähnlich wie bei Datenbanken, Operationen für die Verwaltung und die Auswahl von Daten sowie für den Zugriff auf die Daten. Dies sind elementare Operationen wie:

- SELECT räumlich und attributiv auswählen
- INSERT erzeugen
- UPDATE ändern
- DELETE löschen
- Draw, Scale, Move, …. objektbezogene, geometrische Operationen

(Anmerkung: Datenbanken stellen nur eine Reihe generischer Operationen für die Daten-verwaltung zur Verfügung, da das primäre Ziel die Trennung von Daten und Programmlogik ist.)

Die umfangreichen Operationen und Funktionen, die benötigt werden, um Geodaten auszuwerten und zu analysieren stehen meist als eigene Toolbox zur Verfügung. Diese Operationen müssen objektübergreifend, verschiedene Datenobjekte in Beziehung setzen und verarbeiten können, sie lassen sich daher nicht ohne weiteres einer bestimmten Objektklasse zuordnen.

Beispiel ArcGIS-Toolbox:



Das logische Modell (Zusammenfassung, nach Schilcher)

Im Zuge der logischen Modellierung wird das konzeptionelle Modell in die Datenstrukturen eines Herstellersystems abgebildet

- Durch die Systemauswahl erfolgt die Festlegung auf bestimmte Modellierungstechniken. Beispielsweise werden die Folien- oder Objektklassentechnik zur Strukturierung der Daten, herstellerspezifische topologische Datenstrukturen zur Abbildung der Geometrie, relationale Datenmodelle für die Sachdatenhaltung oder objektorientierte Techniken eingesetzt.
- Die logische Modellierung ist herstellerabhängig!
- Die Abbildung der Geometrie und Topologie erfolgt mit Hilfe der Datenstrukturen, die das dem GIS zugrundeliegende Geometrie- bzw. CAD-Modell beinhaltet. In geringem Umfang lassen sich zu Geometrieelementen (Punkte, Linien, Flächen) auch Attribute abspeichern. Üblicherweise werden jedoch Sachinformationen in Form von Tabellen abgebildet.
- Das entscheidende Merkmal eines integrierten Modells besteht darin, dass Geometrie- und Sachdaten miteinander verbunden sind (Objektbildung). Dies geschieht über entsprechende Zeiger oder Identifikatoren (Geolink).
- Systeme, deren Datenmodellierung objektstrukturiert erfolgt, müssen die Daten nicht zwangsläufig in objektorientierten Datenbanksystemen verwalten.
- Heute ist der GIS-Softwaremarkt dadurch gekennzeichnet, dass eine Vielzahl von Herstellersystemen existieren, die unterschiedlich Modellierungstechniken einsetzen. Die unterschiedlichen Vor- und Nachteile für eine Anwendung müssen bei der Auswahl eines Produktes untersucht werden.
- Die Vielzahl der herstellerspezifischen Datenmodelle verursacht heute noch große Probleme beim Austausch von Geodaten zwischen verschiedenen GIS. Hier können nur durch die Standardisierungs- und Normungsbestrebungen der OGC und ISO/TC211, mit dem Ziel der Interoperabilität, Verbesserungen erzielt werden.

Zusammenfassung: Logisches Konzept für Geoobjekte

Anmerkungen zu Geodatenbanken

Geodatenbanken werden i.d.R. heute auf Basis von Standard DBMS (z.B. Oracle, Informix) implementiert, wobei die besonderen Anforderungen für Geodaten meist über spezielle Datenbankerweiterungen (Extentions) realisiert werden. Beispiel für solche Erweiterungen sind die Spatial-Data-Engine (SDE) der Firma ESRI, die Spatial-Data-Option (SDO) der Firma Oracle oder der Geodatenserver (GDS) der Firma AED-SICAD.

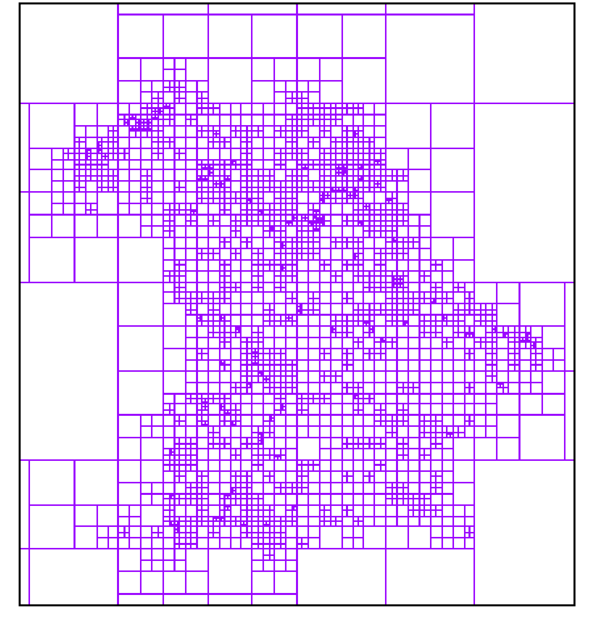
Folgende Merkmale von Geodaten sollten sich in der physischen Struktur niederschlagen

- Gemeinsame, objektstrukturierte Verwaltung von Geometrie- und Sachdaten, Verbindung der verschiedenen Datentypen durch "Geolink"-Relation (spatial link)
- Räumliche Ausdehnung des Datenbestandes in der Datenstruktur berücksichtigen, z.B. durch Partitionierung mit Quadtrees oder R-Bäume
- Räumliche Zugriffe unterstützen, Erweiterung der Abfragesprache um räumliche Operatoren und Funktionen
- Views aus verschiedenen Datenbanken aufgrund des einheitlichen Raumbezugs (Überlagerung)
- Transaktionskonzept für langandauernde Geotransaktionen
- Räumliches oder elementbezogenes Sperrgranulat für Geodaten

Geodaten können sehr umfangreich sein und große Gebiete (Plangebiet) überdecken, sie sind zweidimensional verteilt. Diesem wichtigen Merkmal ist in der DB besonders Rechnung zu tragen. Die Daten werden zweidimensional partitioniert und in Form sogenannter Zellen mittels räumlicher Indizes verwaltet. Den Zellen entsprechen Datenbankblöcke, sie können eine vorgegebene Anzahl von Daten fassen.

In GeoDatenbanken sind für die räumliche Organisation und schnelle Suche in mehrdimensionalen Daten Quadtrees und Rectangle-Trees (R-Bäume mit Variante R*-Baum = optimierter R-Baum) gebräuchlich. Heute wird in DB-Produkten meist der R-Baum oder dessen generalisierte Form GiST als räumlicher Index verwendet.

Quadtree = raumgesteuerter Index → **festes „Universum“** vorgeben

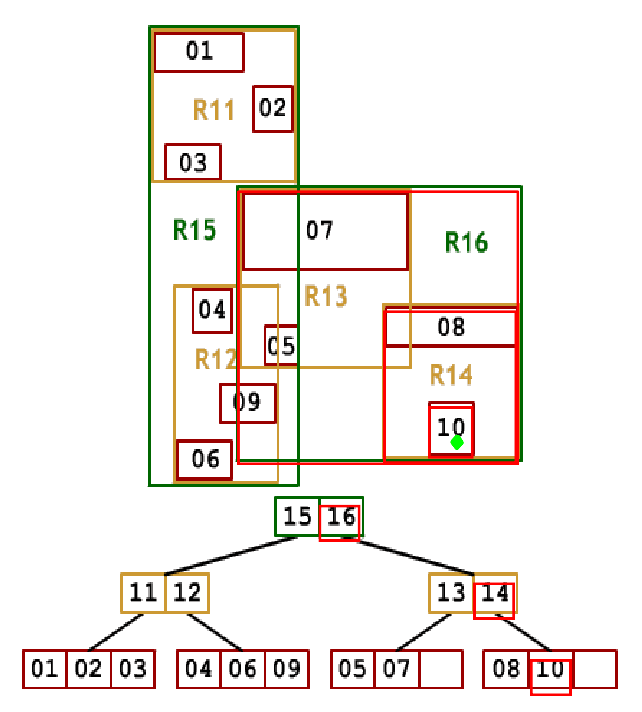


Für die Datenbank ist ein Bereich vorzugeben „Universum“ in dem vorliegende Geodaten gespeichert werden sollen. Über die Festlegung eines rechteckigen Plangebietes und der Definition einer Zelle, die eine vorgegebene Datenmenge aufnehmen kann, wird die räumliche Partitionierung gesteuert. Beim Überlauf einer Zelle wird diese automatisch in vier gleiche Teile aufgeteilt, wobei jede neue Zelle räumlich einem der vier Quadranten der Ausgangszelle entspricht und wieder die vorgegebene Datenmenge aufnehmen kann. Je höher die Datendichte in einem Bereich ist, desto mehr Zellviertelungen ergeben sich. Aus der Plangebietsgröße und der Zellviertelungsstufe lässt sich schnell der geographische Bereich einer Zelle berechnen.

Diese Methode der zweidimensionalen Datenpartitionierung nennt man Quadtree-Struktur, da ein Adressbaum entsteht, der jeweils vier Äste hat. Die Blätter bilden die räumlich zugeordneten Zellen mit den Daten. Die Fragmentierung erfolgt automatisch,

entsprechend der Datendichte in einem Bereich. Die Anzahl der Viertelungen ist zwar begrenzt, aber auch für große Geo-Räume mit hoher Datendichte ausreichend.

Rectangle-Tree = datengesteuerter Index → **„Universum wächst“**



Struktur: balancierter Baum, A. Gutmann

Jede Baumebene gruppiert und aggregiert die Elemente der folgenden Ebene

Ein Objekt wird durch eine Bounding-Box (achsparallele, unregelmäßige Rechtecke) angenähert.

Eigenschaften des R-Baums:

- Alle Knoten haben zwischen m und M Indexeinträge. Wobei $m = \lfloor M/2 \rfloor$ gilt. Die minimale Anzahl m der Indexeinträge darf die Hälfte der Maximalen Anzahl M nicht überschreiten.
- Für jeden Index-Eintrag (I, *tuple-id*) in einem Blatt ist I das kleinste umgebende Rechteck, das das Datenobjekt beinhaltet.
- Jeder Knoten, der kein Blattknoten ist, hat zwischen m und M Söhne. Die Wurzel hat mindestens zwei Söhne.
- Für jeden Eintrag (I, *child-pointer*) in einem Knoten, der kein Blattknoten ist, ist I das kleinste Rechteck, das alle Rechtecke des Kindknotes beinhalten.
- Alle Blätter erscheinen auf derselben Höhe $h = \lceil \log_m N \rceil$ (balanciert), wobei N die Anzahl aller (möglichen) Einträge in den Blattknoten ist.
- Das Splitten eines Knoten wird dann notwendig, wenn die maximale Anzahl M an Indexeinträgen überschritten wird.
- Die Struktur des R-Baumes ist von der Einfügereihenfolge der Daten abhängig, die Rechtecke sind unregelmäßig und können sich stark überlappen.

Durch die räumliche Indexierung der Geodaten wird erreicht, dass bei Zugriffen/Suchen über Koordinaten der Suchaufwand (Performance) der Operationen unabhängig von der Gesamtgröße der Datenbank sind, sondern nur von der Größe des Ausschnittes abhängen. Die räumlichen Indizes können entsprechend auch für 3D-Geometrien erweitert werden.

Folgende Probleme können entstehen, falls keine Datenbank für die Verwaltung von Geodaten verwendet wird, d.h. die Datenverwaltung direkt in einem Dateiverwaltungs-system oder einer sogenannten File-Database erfolgt.

- Beschränkte Zugriffsmöglichkeiten
- Zugriffskonflikte beim Mehrbenutzerbetrieb
- Verlust von Daten
- Redundanz und Inkonsistenz
- Integritätsverletzung, Anomalien
- Sicherheitsprobleme
- Hohe Entwicklungskosten für die Anwenderprogramme

Geotransaktionen

Transaktionen in Geodatenbanken (Geotransaktionen) müssen den ACID-Bedingungen genügen. Sie unterscheiden sich jedoch durch ihre Dauer wesentlich von Transaktionen in herkömmlichen Datenbanken. Während eine Transaktion in rein attributiven Datenbanken üblicherweise im Subsekundenbereich durchgeführt wird, können Geotransaktionen Stunden oder sogar Tage dauern. Die Nebenläufigkeit von Transaktionen (Parallelzugriffe) für lesende Zugriffe muss dabei ohne Einschränkung erhalten bleiben.

Damit auch ändernde Zugriffe zu möglichst wenigen Kollisionen in der Datenbank führen, muss das Transaktionsgebiet und damit auch die Sperrung betroffener Datenbereiche minimiert werden. Das Sperrgranulat in der SICAD-GDB ist die Zelle, oder das Geometrieelement. Objektbezogene Änderungen sollten deshalb auf der Ebene minimaler Mengen von Geoobjekten – die hier als Minimalkollektiv bezeichnet werden – durchgeführt werden. Unter einem Minimalkollektiv wird dabei die kleinstmögliche Menge von Teilobjekten verstanden, die für eine Geotransaktion mit änderndem Zugriff erforderlich ist.

Operationen mit Minimalkollektiven können auf die elementaren Fälle Punkt-, Linien- und Flächenänderung zurückgeführt werden. Damit Geotransaktionen mit änderndem Zugriff von möglichst kurzer Dauer sind, und sie nur kleine Bereiche der Geodatenbank auf Ebene der Zellen sperren, sind Minimalkollektive für die Durchführung einer Änderung zu verwenden. Die Auswahl der Ausgangselemente, die von einer Änderung betroffen sind, kann dazu geometrisch in einem Übersichtsbild oder attributiv über die Organisationsattribute erfolgen.

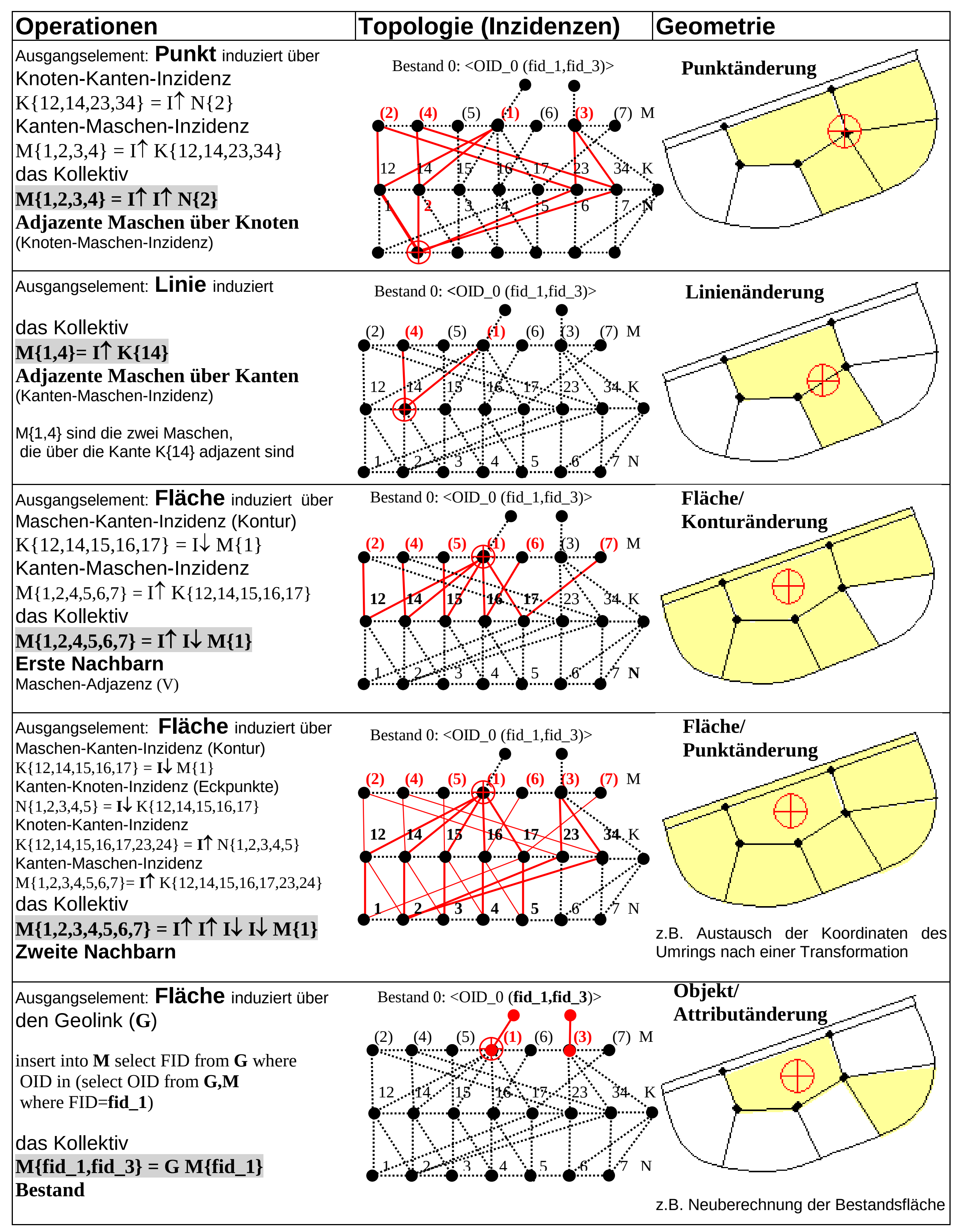
Die Inzidenzoperatoren sind mengenorientiert, daher kann die Ausgangsmenge auch aus mehreren zu ändernden Elementen vom Typ Punkt, Linie oder Fläche bestehen. Durch die Navigation mit den Inzidenzoperatoren entsprechend der Grundmuster erhält man dann die minimale zugehörige Menge von Objekten bzw. Objektteilen, die von der Änderung betroffen sind.

Allgemeine Zugriffe erfolgen i.d.R. für Auskunftszwecke, d.h. nur „Lesend“, dabei wird durch das Transaktionskonzept gewährleistet, dass keine Behinderung durch andere lesende oder ändernde Zugriffe entstehen. Man kann generell zwischen objektbezogenen und wertebezogenen Zugriffen unterscheiden, die mittels generischer Objektoperationen für das direkte Auffinden oder generischer Werteoperationen für Attribute des Objekts gemäß seinen Werteaufbau realisiert werden. Für topologische objektbezogene Zugriffe können die Inzidenzbeziehungen genutzt werden, für semantische objektbezogene Zugriffe die Attribute.

Die optimale physische Konfiguration und Parametrisierung der Datenbank, in Abhängigkeit von der Datenstruktur und Benutzerorganisation der Fachanwendung, ist eine wichtige Aufgabe des Systemkonzepts, da nur dadurch die Datenkonsistenz und Datensicherheit gewährleistet ist, aber auch eine gute Performance und eine wirtschaftliche Ressourcenauslastung (Hauptspeicher, Platten) erreicht werden kann.

Anmerkung zur OID: Aus fachlicher Sicht ist die OID, insbesondere wenn sie durch ein allgemeines Surrogat gebildet wurde, kein geeignetes Zugriffskriterium, da sie ein Objekt zwar eindeutig bezeichnet, das Objekt aber anonym lässt. Die OID kann man mit der Personalausweisnummer vergleichen, die zwar eine Person eindeutige identifiziert, über die man sie aber nicht ansprechen kann. Für den semantischen objektbezogenen Zugriff benötigt ein Objekt deshalb zusätzlich Personalien („Namen, Geburtsdatum“), also eine fachlich definierte, eindeutige Bezeichnung.

Grundlegende Update-Fälle für Geoobjekte mit topologischen Beziehungen



Operationen	Topologie (Inzidenzen)	Geometrie
Ausgangselement: **Punkt** induziert über Knoten-Kanten-Inzidenz K{12,14,23,34} = I↑ N{2} Kanten-Maschen-Inzidenz M{1,2,3,4} = I↑ K{12,14,23,34} das Kollektiv **M{1,2,3,4} = I↑ I↑ N{2}** **Adjazente Maschen über Knoten** (Knoten-Maschen-Inzidenz)	Bestand 0: <OID_0 (fid_1,fid_3)>	**Punktänderung**
Ausgangselement: **Linie** induziert das Kollektiv **M{1,4}= I↑ K{14}** **Adjazente Maschen über Kanten** (Kanten-Maschen-Inzidenz) M{1,4} sind die zwei Maschen, die über die Kante K{14} adjazent sind	Bestand 0: <OID_0 (fid_1,fid_3)>	**Linienänderung**
Ausgangselement: **Fläche** induziert über Maschen-Kanten-Inzidenz (Kontur) K{12,14,15,16,17} = I↓ M{1} Kanten-Maschen-Inzidenz M{1,2,4,5,6,7} = I↑ K{12,14,15,16,17} das Kollektiv **M{1,2,4,5,6,7} = I↑ I↓ M{1}** **Erste Nachbarn** Maschen-Adjazenz (V)	Bestand 0: <OID_0 (fid_1,fid_3)>	**Fläche/ Konturänderung**
Ausgangselement: **Fläche** induziert über Maschen-Kanten-Inzidenz (Kontur) K{12,14,15,16,17} = I↓ M{1} Kanten-Knoten-Inzidenz (Eckpunkte) N{1,2,3,4,5} = I↓ K{12,14,15,16,17} Knoten-Kanten-Inzidenz K{12,14,15,16,17,23,24} = I↑ N{1,2,3,4,5} Kanten-Maschen-Inzidenz M{1,2,3,4,5,6,7}= I↑ K{12,14,15,16,17,23,24} das Kollektiv **M{1,2,3,4,5,6,7} = I↑ I↑ I↓ I↓ M{1}** **Zweite Nachbarn**	Bestand 0: <OID_0 (fid_1,fid_3)>	**Fläche/ Punktänderung** z.B. Austausch der Koordinaten des Umrings nach einer Transformation
Ausgangselement: **Fläche** induziert über den Geolink (G) insert into **M** select FID from **G** where OID in (select OID from **G,M** where FID=**fid_1**) das Kollektiv **M{fid_1,fid_3} = G M{fid_1}** **Bestand**	Bestand 0: <OID_0 (**fid_1,fid_3**)>	**Objekt/ Attributänderung** z.B. Neuberechnung der Bestandsfläche

Die angegeben Operationen sind wie Zuweisungen in einem Programm zu lesen, der linken Seite wird das Ergebnis des Ausdrucks auf der rechten Seite zugewiesen.

Anmerkung: Der Rand des Minimalkollektivs kapselt die Änderung, das optimale Sperrgranulat sind die Geometrieelemente.

Modellierung im GIS
(Zusammenfassung, Schilcher, TUM)

Worauf es bei der Modellierung ankommt !

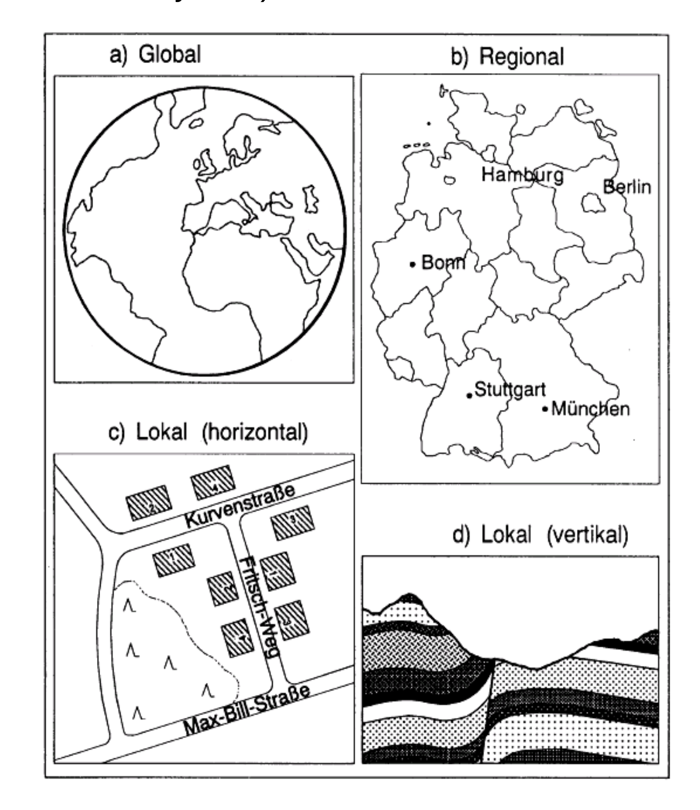
- Das Datenmodell bestimmt die Leistungsfähigkeit eines GIS. Es bildet ein wichtiges Entscheidungskriterium bei der Auswahl eines Herstellersystems
- Das Datenmodell bestimmt den Aufwand und die Kosten für Erfassung und Pflege eines GIS. Die Voraussetzungen für ein wirtschaftlich erfolgreiches Projekt werden damit bereits mit der Modellierung geschaffen.
- Es gibt heute keine universell einsetzbaren Datenmodelle, die allen Anwendungen gerecht werden und optimale Lösungen garantieren. In der Praxis dominieren projektspezifische, anwendungs- und herstellerabhängige (propritäre) Datenmodelle.
- In Fach-GIS gibt es einen Trend zu einer Art "Vormodellierung" in Form von Spartenlösungen und Fachschalenkonzepten durch Hersteller, Fachverbände oder Anwender. Damit kann der Modellierungsaufwand verringert und die Einführungszeit verkürzt werden.
- Am meisten verbreitet sind heute relationale Modelle in Verbindung mit RDBMS. Objektorientierte Datenmodelle mit RDBMS bzw. ORDBMS sind die Lösungen der Zukunft.
- Es gibt keine universell nutzbaren Geodaten und es gibt keine maßstabslosen Geodaten. Bei der Modellierung muss zwischen maßstäblichen (lagetreuen, nichtgeneralisierten) und generalisierten Daten unterschieden werden.
- Die wichtigsten Qualitätskriterien sind die Aktualität und Konsistenz der Daten (nicht die Genauigkeit). Die Aktualisierung muss bereits im Datenmodell berücksichtigt werden. Falls dieser Aspekt vernachlässigt wird, führt er häufig zu "Datenfriedhöfen" oder hohen Migrationskosten bei Systemveränderungen.
- Es gibt keine universell einsetzbaren Herstellersysteme. Es gibt einen deutlichen Trend zu sogenannten "Produktfamilien" von GIS-Anbietern (Komponenten-technologie) mit einer abgestuften, an den Aufgaben des jeweiligen Endnutzers orientierten Funktionalität.
- In Zukunft müssen bei der Modellierung stärker als in der Vergangenheit bereits vorhandene Geoinformationssysteme berücksichtigt werden. Das gleiche gilt für die Einbeziehung von Standards (OGC) und Normen (ISO).
- Bei großen GIS-Projekten sollte die Modellierung in Stufen geplant werden mit iterativer Verfeinerung und Korrektur des Datenmodells (Evolution).
- **Modellierungs-know-how zählt zu den Schlüsselqualifikationen der GIS-Mitarbeiter und sollte in Unternehmen aufgebaut werden.** Ein ausschließlich externer Aufbau bzw. Einkauf ist problematisch, da gerade bei der Modellierung mit laufenden Anpassungen und Weiterentwicklungen zu rechnen ist (Evolution). Dieses know-how ist nicht nur für den Neuentwurf eines Modells erforderlich, sondern auch zur qualifizierten Nutzung eines bestehenden Modells.
- Es gibt erhebliche Defizite in der Standardisierung und Normierung von GIS. OGC und ISO werden Verbesserungen bringen, zu welchem Zeitpunkt praxisreife Lösungen verfügbar sein werden, ist gegenwärtig noch nicht abzusehen.

6. Raumbezug, Geokodierung

In einem GIS werden georäumliche Erscheinungen modelliert, d.h. es werden Teile der Erdoberfläche nach Lage und Höhe abgebildet. Grundlage für diese Abbildung ist ein definiertes **Raumbezugssystem** (Spatial Reference System), das durch

- ein geodätisches Bezugssystem
- ein zugeordnetes Koordinatensystem
- geeignete Elemente, die das Bezugssystem im GIS repräsentieren (Bezugsrahmen, Anschlusspunkte)

festgelegt wird.



In Abhängigkeit von der Auflösung der Geodaten und von der räumlichen Ausdehnung (Größe und Dimension) des Gebietes, für das ein GIS angelegt werden soll, muss das Raumbezugssystem geeignet sein lokale, regionale oder globale Abbildungen nach Lage und Höhe zu ermöglichen.

Der einheitliche Raumbezug ist das verbindende Element ("Basis-Join") im GIS, das die räumliche Zuordnung und fachübergreifende Verknüpfung der Geodaten ermöglicht. Er ist das wesentliche Charakteristikum von Geodaten.

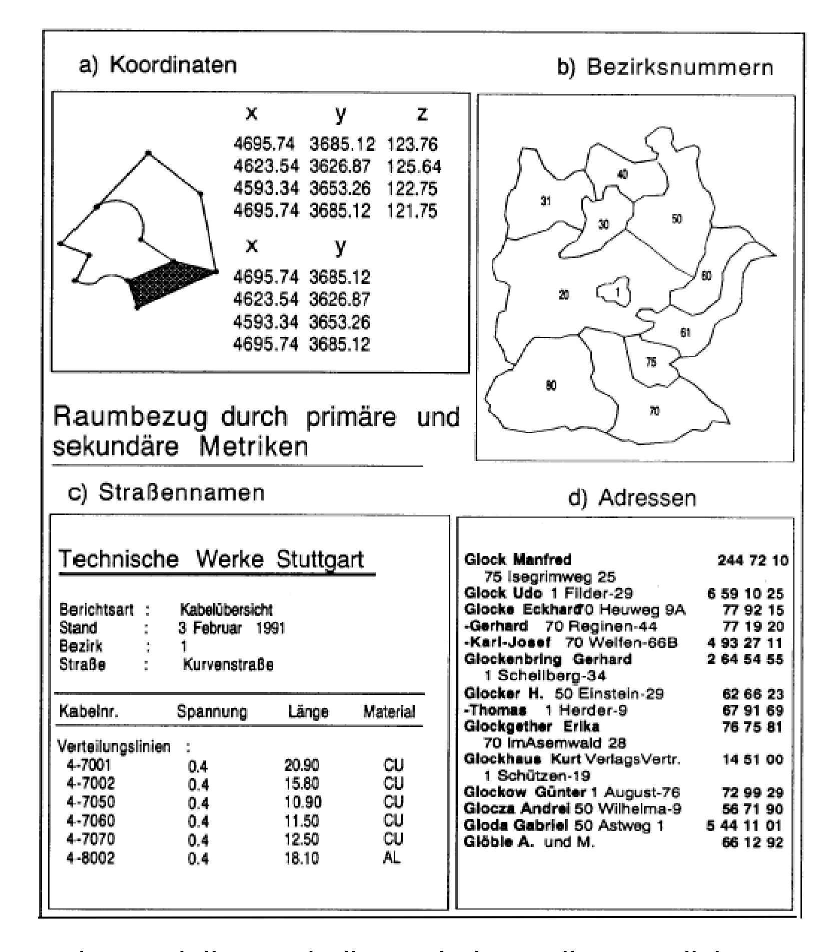
Fragen in denen das Wo?, das Wie groß?, das Wie weit bis? eine Rolle spielen, können nur beantwortet werden, wenn den Geodaten ein definiertes Bezugssystem zugrunde liegt, das die räumliche Verbindung zwischen Modell und dem GeoRaum darstellt.

Die theoretische Festlegung des Raumbezugs erfolgt durch die Wahl eines geodätischen Bezugssystems, dem für die Zahlengebung, dem Geokode, ein geeignetes Koordinatensystem zugeordnet wird. Geodätische Bezugssysteme werden mittels physikalischer und geometrischer Festlegungen beschrieben, die definieren wie die Form des Erdkörpers approximiert und wie der Bezug zu einer geeigneten Referenzfläche hergestellt wird.

Die Nutzung eines Bezugssystem (Spatial Reference System) für den Raumbezug in einem GIS oder für geodätische Zwecke erfordert seine „Materialisierung“, den sogenannten Referenzrahmen. Referenzrahmen (Frame) werden üblicherweise durch Netze diskreter Anschlusspunkte realisiert, dies können abgemarkte Punkte im Feld oder virtuelle Punkte in einer Datenbank sein. Diese Verbindungselemente müssen für den Anschluss an das Raumbezugssystem, d.h. zur Geokodierung der Objekte - für die Ermittlung der benötigten Koordinaten - hinreichend genau sein. Die Anschlusspunkte können z.B. einen Basisgeodatenbestand (z.B. DFK, AFIS) entnommen oder durch die Abbildung eines Koordinatengitters (Flurkartenblattecken) realisiert werden oder sie ergeben sich direkt aus Messungen (z.B. GPS).

Mit den Koordinaten wird in einem Raumbezugssystem festgelegt wo sich ein Geoobjekt auf der Erdoberfläche befindet, seine Lage, und welche Form es hat.

Geokodierung:
Man spricht von der Geokodierung, wenn für jeden einzelnen Punkt bzw. Stützpunkt der Geodaten eine Koordinate im Raumbezugssystem ermittelt wurde.



Georeferenzierung:
Bilddaten werden häufig nur über einige Punkte mittels Transformationen in das Raumbezugssystem eingepasst, wobei die Bildkoordinaten selbst unverändert bleiben, man bezeichnet das Bild dann als georeferenziert.

Direkter/indirekter Raumbezug:
Raumbezug kann gegeben sein direkt über Koordinaten (primär) oder indirekt über Tabellen, z.B. Adressen, Postleitzahlen oder Flurstücksnummern denen Koordinaten zugeordnet sind (sekundär). Daten mit einer sekundärer Metrik werden als georeferenzierte Fachdaten bezeichnet, da sie ihren Raumbezug indirekt durch die semantische Verknüpfung mit speziellen Tabellen erhalten, die räumliche Zuordnungen sematischer Daten mittels entsprechender Koordinaten enthalten.

Geodätische Bezugssysteme

Aus der Figur der Erde heraus, die in guter Näherung ein an den Polen abgeplattetes Rotationsellipsoid ist, ergibt sich für die Abbildung der Erdoberfläche ein dreidimensionales Problem. Einem GIS könnte zwar auch direkt ein dreidimensionales geozentrisches Koordinatensystem zugrunde gelegt werden, da jedoch

die Karte die wichtigste Darstellungsform von Geodaten ist,
werden in GIS topozentrische, flächenbezogene Lagekoordinaten bevorzugt,
die durch geeignete Höhenangaben ergänzt werden.

Dazu müssen die abzubildenden Geoobjekte auf eine mathematisch beherrschbare Referenzfläche bezogen werden, deren Lage zum Erdkörper (geodätisches Datum) festgelegt ist und auf der ein Koordinatensystem ausgebreitet werden kann.

Referenzflächen

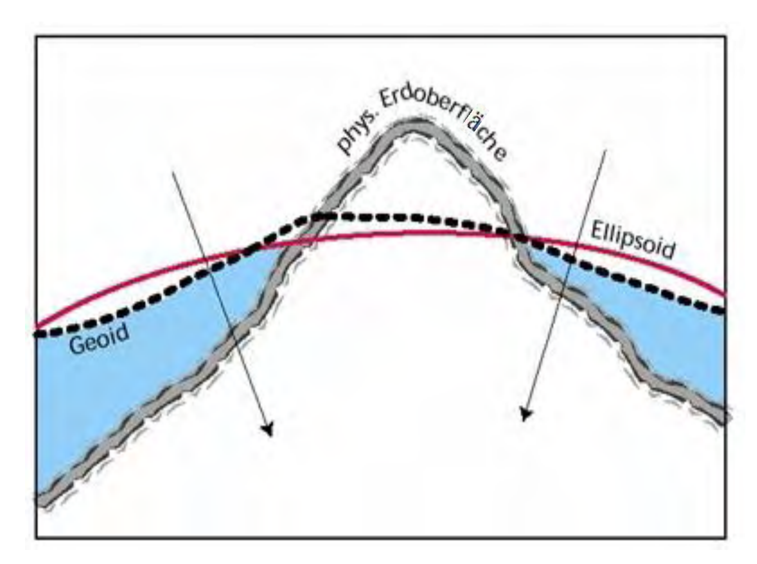
Das Erdschwerefeld stellt die Hauptorientierung (Lotrichtung) für exakte Messungen dar, es wäre damit naheliegend, die Referenzfläche physikalisch zu definieren. In der Geodäsie wurde dazu das Geoid eingeführt. Das Geoid ist eine Äquipotentialfläche des Schwerefeldes in Höhe eines mittleren Meeresspiegels, d.h. eine Fläche mit konstantem Schwerepotential; auf einer solchen Fläche würde kein Wasser fließen. Das Geoid hat jedoch eine sehr komplizierte Form, die nur durch Reihenentwicklungen approximiert werden kann, damit ist es zur Definition von Lagekoordinaten nicht geeignet. Referenzflächen werden daher mathematisch definiert und mit dem Erdkörper in Beziehung gesetzt.

Je nach Aufgabenstellung und Genauigkeitsanforderung gibt es verschiedene Möglichkeiten die Erdfigur durch eine Referenzfläche zu approximieren.

Lokale Vermessung	**< 10 km**	**Ebene**	**kartesische Koordinaten**
Landesvermessung	**< 200 km**	**Kugel**	**sphärische Koordinaten**
	> 200 km	**Rotationsellipsoid**	**geodätische Koordinaten**
Geodäsie	**global**	**Geozentrum**	**3D-kartesische Koordinaten**
		Geoid	**Lotrichtung, Lotabweichung**
Geographie	**Länder, Welt**	**Kugel**	**geographische Koordinaten**

Für kleinmaßstäbliche Anwendungen (<1:1Mio), wie sie in der Geographie üblich sind, kann die Erde als Kugel betrachtet werden.

Bei großmaßstäblichen Anwendungen können für kleine Gebiete direkt ebene kartesische Koordinaten verwendet werden (5 - 10 km Ausdehnung). Für größere Gebiete wie z.B. das Land Bayern sind die Grundlagen einer Landesvermessung erforderlich, da sich sonst die Einflüsse der Erdkrümmung als Modellfehler auswirken.



Den heutigen Landesvermessungen liegen meist ellipsoidische Referenzflächen - Rotationsellipsoide - zugrunde, in Bayern das "Besselellipsoid".

Parameter	**Bessel**	**Hayford (ED50)**	**Krassowski**	**GRS80 (GPS)**
Große Halbachse ($a_{[m]}$)	6 377 397.155	6 378 388.000	6 378 245.000	6 378 137.000
Kleine Halbachse ($b_{[m]}$)	6 356 078.963	6 356 911.946	6 356 863.019	6 356 752.314
Abplattung (f)	1:299.1528128	1:297	1:298.3	1:298.2572236
Polkrümmung ($c_{[m]}$)	6 398 786.849	6 399 936.608	6 399 698.902	6 399 593.626

Auf den Referenzellipsoiden wurden durch die Landesvermessungen Koordinatensysteme ausgebreitet, wozu früher hierarchische Festpunktnetze durch Winkel- und Streckenmessungen bestimmt worden sind. Solche Punktfelder bezeichnet man heute auch als Benutzersegment oder Koordinatenframe eines Bezugssystems, da sie das Koordinatensystem physisch realisieren, so dass lokale Messungen einfach an das landesweite oder globale Koordinatensystem angeschlossen werden können.

Heute werden globale, geozentrische Bezugssysteme durch die GNS-Systeme realisiert, in denen die Bahndaten der zur Positionierung verwendeten Satelliten gegeben sind (z.B. WGS84, ETRS89).

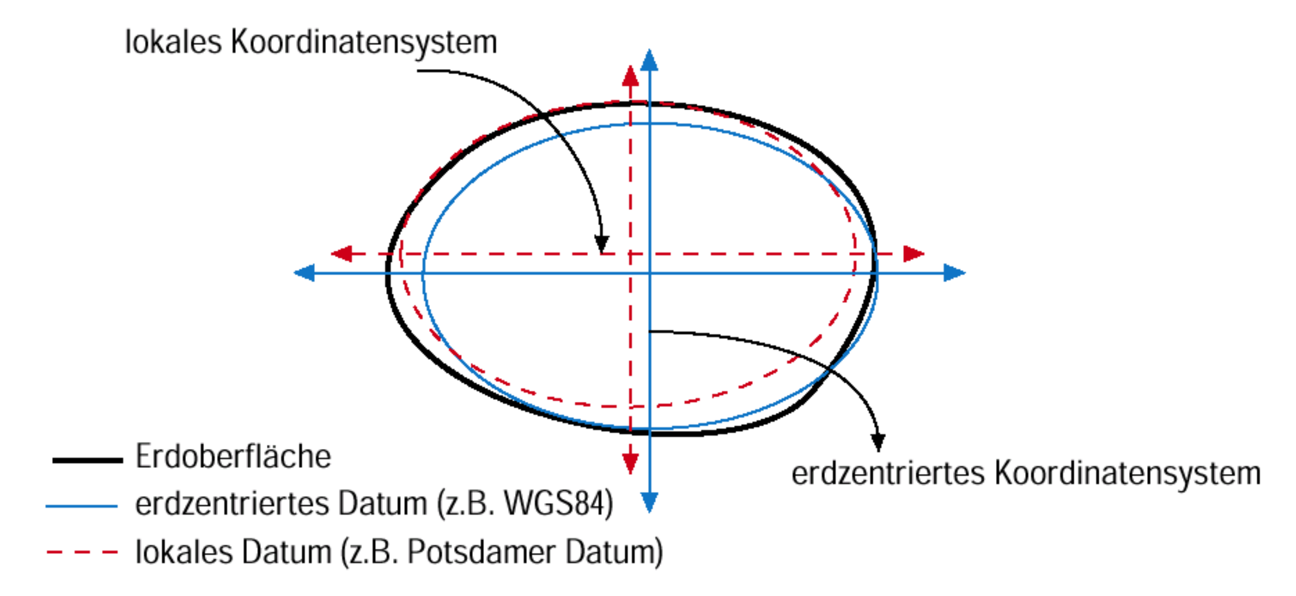
Alle Messungen müssen für die Berechnung der Koordinaten auf die Referenzfläche reduziert werden, um Modellfehler zu vermeiden. Reduktionen können für topographische Messungen meist vernachlässigt werden, da sie unter der erforderlichen Messgenauigkeit liegen.

Geodätisches Datum (klassisch)

Der räumliche Bezug einer Referenzfläche zum Erdkörper wird durch das geodätische Datum hergestellt, wobei man für seine geometrische Definition im Wesentlichen über sechs Freiheitsgrade, drei Translationen und drei Rotationen, verfügen muss. Die heute verwendeten Landesvermessungen entstanden meist zu Beginn des 19. Jahrhunderts. Zur Bestimmung des geodätischen Datums wurden i.d.R. für den Fundamentalpunkt eines Netzes die Beziehungen zwischen seinen physikalisch bestimmten Koordinaten $P_0(\varphi_0,\lambda_0,H_0)$ und seinen geometrisch definierten Koordinaten $P_0(B_0,L_0,h_0)$, bezogen auf das gewählte Referenzellipsoid (a,b), festgelegt. Lage- und Höhe wurden getrennt betrachtet. Für die Lagenetze sind üblicherweise die Festlegungen $B_0 = \varphi_0$ und $L_0 = \lambda_0$ getroffen worden, d.h. die geodätischen Koordinaten (B_0,L_0) des Fundamentalpunktes wurden gleich seinen astrogeodätischen Koordinaten (φ_0,λ_0) gesetzt und damit die Lotabweichungen im Fundamentalpunkt $\xi_0 = \eta_0 = 0$. Als Bezugsfläche wurde im allgemeinen ein bestanschließendes Ellipsoid verwendet.

Wenn die kleine Halbachse b des Rotationsellipsoids (a,b) parallel zur Erdachse steht, gelten folgende Beziehungen:	B_0,L_0	**geodätische Koordinaten**
$B_0 = \varphi_0 - \xi_0$	A_0	**geodätisches Azimut**
$L_0 = \lambda_0 - \eta_0/\cos B_0$	φ_0,λ_0	**Lotrichtung (astrogeod. Koord.)**
$A_0 = a_0 - \eta_0/\tan B_0$	a_0	**astronomisches Azimut**
$H_0 = h_0 - N_0$	ξ_0, η_0	**Lotabweichungen**
	h_0	**Ellipsoidhöhe**
	H_0	**Meereshöhe**
	N_0	**Geoidundulation bezogen auf (a,b)**

Durch diese Festlegungen fällt im Fundamentalpunkt die Ellipsoidnormale mit der Lotrichtung zusammen. Für die Orientierung der Lagenetze wurde das geodätische Azimut (A_0) gleich dem astronomischen Azimut (a_0) gesetzt, wodurch die kleine Halbachse (b) des Rotationsellipsoids (a,b) näherungsweise parallel zur Erdachse gestellt wird. Schließlich gilt für die ellipsoidische Höhe (H_0) des Fundamentalpunktes und seiner Meereshöhe (h_0) noch die Beziehung $h_0 = H_0+N_0$, wobei die Geoidundulation (N_0) den Höhenunterschied zwischen dem Referenzellipsoid und Geoid angibt.



Ein Rotationsellipsoid ist durch die Angabe von zwei Parametern (a,f oder a,b) geometrisch bestimmt. In der Landesvermessung unterscheidet man mittlere (erdzentrierte) und bestanschließende (lokale) Ellipsoide.

Bis vor einigen Jahrzehnten wurde die Bestimmung der Referenzflächen weitgehend aus astrogeodätischen Lotabweichungen durchgeführt, die nur auf den Kontinenten gemessen werden können. Solche Beobachtungen liefern bestanschließende Ellipsoide, die sich in ihren Dimensionen und ihrer Lagerung der physikalisch definierten Erdoberfläche, dem Geoid, im jeweiligen Bereich (lokal) der Landesvermessung optimal anpassen. Ein mittleres Ellipsoid und eine geozentrische (erdzentrierte) Lagerung der Systeme waren damit nicht erreichbar, diese konnten erst mit Hilfe von Satellitenmessungen bestimmt werden. Bei modernen Bezugssystemen wird die Lagerung durch die räumlichen Koordinaten von ausgewählten Referenzstationen festgelegt.

Koordinatensysteme (CRS)

Für ein Raumbezugssystem ist es erforderlich, dem geodätischen Bezugssystem ein geeignetes Koordinatensystem zuzuordnen und durch Anschlusspunkte zu realisieren.

Die Abbildung der Erde in GIS erfolgt heute noch meist nach Lage (2D-Grundriss) und Höhe (1D-Relief) getrennt, die wichtigsten Gründe dafür sind:

- in Abhängigkeit von der Anwendung wird der Lage- und Höheninformation eine unterschiedliche Bedeutung beigemessen, dies hat Auswirkungen auf die nötige Datendichte und die Aufnahmegenauigkeit der verschiedenen Dimensionen
- die topozentrische Anschaulichkeit erfordert flächenbezogene Koordinaten
- terrestrische Lage- und Höhenmessungen beruhen auf verschiedenen Messverfahren
- es werden unterschiedliche Referenzflächen verwendet, für die Lage ein Rotationsellipsoid, für die Höhe das Geoid oder Quasigeoid (Meereshöhen).
- die vollständige 3D-Erfassung und Modellierung eines großen Gebietes ist sehr aufwendig

Moderne Messverfahren (Satellitennavigation oder Fernerkundung) liefern zunehmend direkt 3D-Daten, so dass die dreidimensionale Modellierung von Geoobjekten immer mehr an Bedeutung gewinnt.

Koordinatensysteme		Koordinatentyp	Bezugssystem	
Lage	Amtl. Koord. Bayern	eben/konform, GK	DHDN, Rauenberg	Bessel
	seit 2006 Umstellung	eben/konform, UTM	ETRS89	GRS80
	UTM-Koordinaten	eben/konform, UTM	ED50	Hayford
	seit 1996		WGS84	GRS80
Histo-risch	Bayer. Soldner-Koord.	sphärisch	Frauenkirche München	Laplace, Kugel
	Amtl. Koord. DDR	eben/konform, GK	S42/83, Pulkowo / St. Petersburg (EAGN)	Krassowsky
Höhe	Normalhöhen (Meereshöhen)	Lotlinie	DHHN, Pegel Amsterdam (NHN)	Quasigeoid
	Ellipsoidhöhen	Lot auf Ellipsoid	jeweiliges Referenzellipsoid	
Raum	GPS-Koordinaten	3D-kartesisch	WGS84/ETRS89	Geozentrum

Für die landesweite Realisierung eines geodätischen Bezugssystems wurden Punktnetze verschiedener Ordnung gemessen, die hierarchisch aufgebaut sind. Die Netze sind soweit verdichtet worden, dass lokale Messungen einfach angeschlossen werden können. Für lokale Messungen sollen i.d.R. ebene Rechnungen möglich sein und nur geringe Reduktionen anfallen.

Einteilung Geodätischer Bezugssysteme, nach *Schnädelbach (1996)*

Zur Definition Geodätischer Bezugssysteme gibt es folgende Möglichkeiten:

Dreidimensionale kartesische geozentrische Koordinatensysteme $P(X_g,Y_g,Z_g)$

Der Ursprung liegt im Massenmittelpunkt der Erde (g=Geozentrum), die Z_g-Achse geht in Richtung einer mittleren Rotationsachse der Erde, die X_g-Achse bildet die Schnittlinie eines Bezugsmeridians mit dem Äquator. In solchen Systemen werden künstliche Erdsatelliten koordiniert, z.B. des GPS im WGS84. Es eignet sich schlecht für die anschauliche Darstellung von Geoinformation und Vermessungsergebnissen.

Ellipsoidische Bezugssysteme

Die Erdoberfläche lässt sich in guter Näherung durch ein an den Polen abgeplattetes Rotationsellipsoid approximieren, dessen Bezug zum Erdkörper durch das geodätische Datum festgelegt wird. Damit entstehen mit dieser Fläche verbundene, also geometrisch definierte ellipsoidische Systeme. Die Koordinaten lassen sich je nach Erfordernis in verschiedenen Formen darstellen.

- **Dreidimensionale kartesische Koordinaten P(X,Y,Z):** Der Ursprung liegt im Figurenmittelpunkt des Ellipsoids, die Z-Achse ist identisch mit der Rotationsachse, die X-Achse wird von der Schnittlinie eines Bezugsmeridians mit dem Äquator gebildet. Diese Systeme eignen sich ebenfalls schlecht für anschauliche Darstellungen von Vermessungsergebnissen und damit auch für GIS, sind aber wichtig für die Transformation zwischen verschiedenen Bezugssystemen.

- **Geodätische Koordinaten P(B,L,h):** Von Vorteil ist bei diesen Systemen der enge Bezug zur Lotrichtung und damit zu den Begriffen „Höhe“ und „Grundriss“ sowie die Möglichkeit der einheitlichen Koordinierung auch der großräumigsten Projekte. Vermessungsarbeiten können gut dargestellt werden, nachteilig sind die schwierigen Berechnungen bei großmaßstäblichen Arbeiten, da keine ebene Trigonometrie verwendet werden kann. GIS-Systeme erlauben heute ihre Verwendung für die Speicherung von Geodaten, für die Analyse und Präsentation können die geodätischen Koordinaten dann „on the fly“ projiziert werden.

- **Projektionskoordinaten $P(y,x,L_0)$:**
  Mit diesen Koordinaten werden Teile der Ellipsoidoberfläche durch kartesische Koordinaten beschrieben, so dass bei allen Folgearbeiten die ebene Trigonometrie verwendet werden kann. Das System wird in der dritten Dimension durch Ellipsoid- oder Meereshöhen ergänzt.

  Nachteilig ist wegen der auftretenden Verzerrungen die begrenzte Größe des Anwendungsgebietes. Die Verzerrungen können jedoch mit entsprechenden Korrekturgliedern, je nach Genauigkeitsansprüchen ermittelt werden.

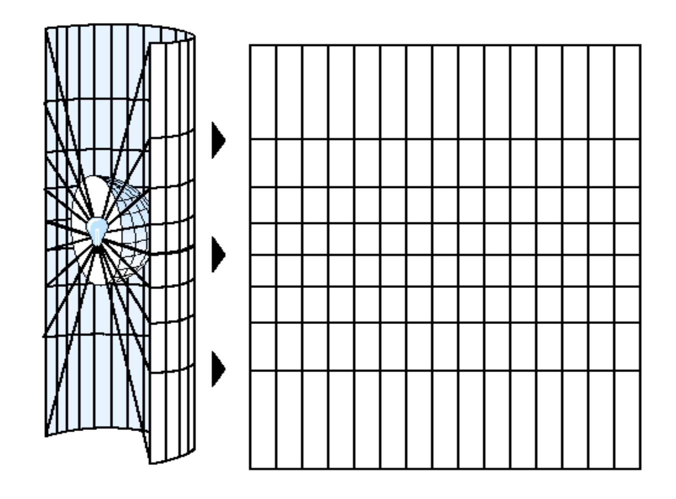
  Die Anwendung von Projektionskoordinaten ist die Regel für den Raumbezug in GIS-Projekten und für lokale Vermessungen *(siehe Vorlesung Geodätische Bezugssysteme)*.

P(X,Y,Z) ↔ P(B,L,h) ↔ P(y,x,H).

Wichtige Typen von Projektionskoordinaten sind:

Gauß- Krüger-Koordinaten (eben, konform) oder

Soldnerkoordinaten (rechtwinklig-ellipsoidisch oder rechtwinklig-sphärisch).



- **Gebrauchskoordinaten P(RW,HW):**

  Um die Anwendung von Projektionskoordinaten zu vereinfachen, kann man sie zu so genannten Gebrauchskoordinaten modifizieren.

  Dabei werden zum Beispiel Konstanten addiert, damit keine negativen Koordinatenwerte entstehen oder ein mittlerer Maßstab eingeführt, um die Beträge der Projektionsverzerrungen zu verringern. So sind z.B. Gauß-Krüger- und UTM-Koordinaten von der Projektion her ebene konforme Koordinaten, die jedoch aus pragmatischen Gründen unterschiedlich zu Gebrauchskoordinaten modifiziert werden.

Gauß-Krüger-Koordinaten

Die wichtigste Projektion im Deutschen Vermessungswesen ist die ebene konforme Abbildung nach C. F. Gauß. Diese Abbildung hat Gauß im Rahmen der Hannoverschen Landesvermessung zwischen 1820 und 1830 entwickelt. Das Verfahren wurde später von Krüger überarbeitet und dokumentiert und wird deshalb als Gauß-Krüger-Abbildung bezeichnet. Sie liegt heute noch dem amtlichen Bayerischen Koordinatensystem zugrunde und wird auch als Projektion für die großmaßstäblichen topographischen Karten verwendet. Bei der Gauß-Krüger-Abbildung werden die geodätischen Koordinaten (B,L) auf einem Ellipsoid konform (winkeltreu im Differenziellen) in ein ebenes kartesisches Koordinatensystem (x,y) abgebildet, wobei:

- der Schnittpunkt des durch das abzubildende Gebiet verlaufenden Bezugsmeridians (Hauptmeridian) mit dem Äquator die ebenen Koordinaten (x,y) = (0,0) erhält
- das Bild des Hauptmeridians eine Gerade ergibt, die die Abszissenachse (x) des Systems bildet (gegenseitige Festlegung der Orientierung der beiden Koordinatensysteme)
- die Länge auf der Abszissenachse zum entsprechenden Bogen des Hauptmeridians in einem konstanten Verhältnis steht (gegenseitige Festlegung des Maßstabes)

Die Abbildungsgleichungen ergeben sich aus der konformen Abbildung

$$\boldsymbol{x} + \boldsymbol{iy} = \boldsymbol{F}(\boldsymbol{q} + \boldsymbol{il}) \quad \text{mit} \quad \boldsymbol{F}'(\boldsymbol{q} + \boldsymbol{il}) = \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{l}} - \boldsymbol{i}\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{l}}$$

$\boldsymbol{q}$	*isometrische Breite = artanh(sinB) – e artanh(e sinB*
$\boldsymbol{l = L - L_0}$	*Längenunterschied zum Hauptmeridian (L_0)*
$\boldsymbol{B,L}$	*geodätische Breite und Länge*

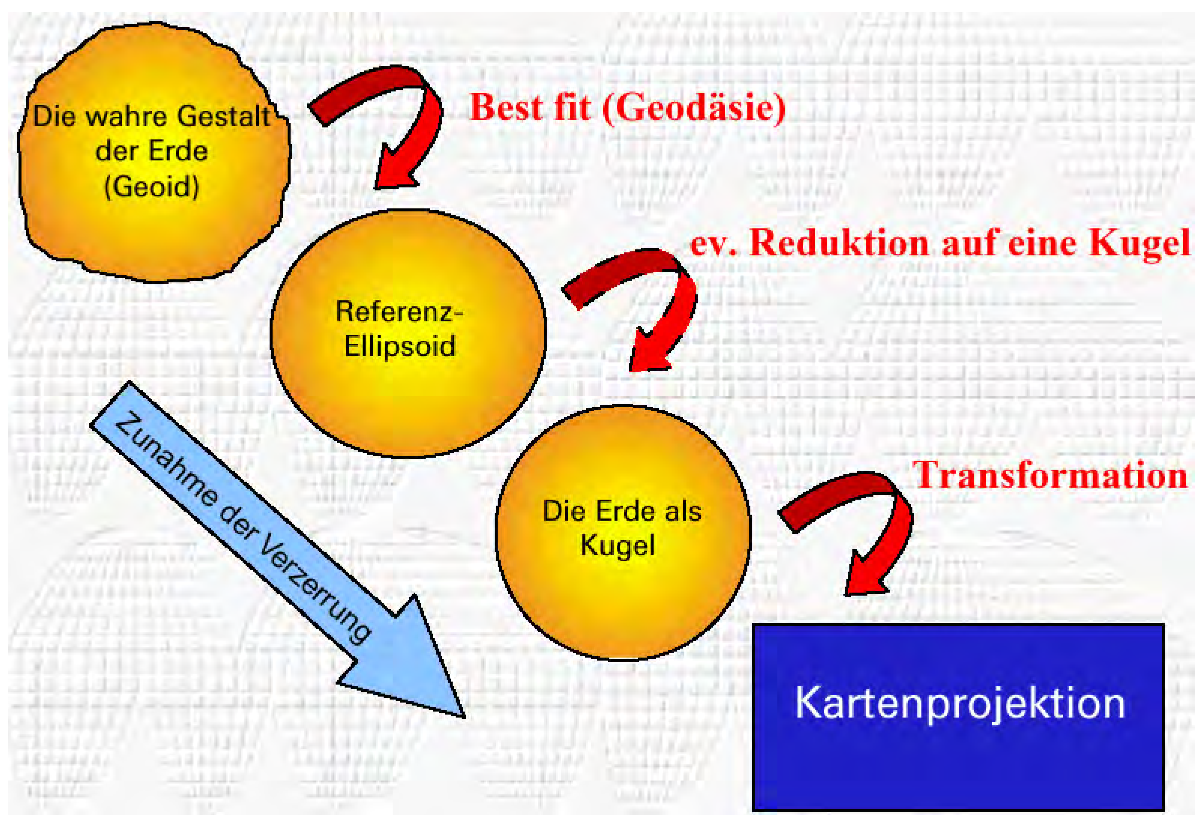
Diese Funktion kann geschlossen nur mit einem elliptischen Integral dargestellt werden, deshalb müssen die Abbildungsgleichungen in Reihen entwickelt werden.

Wird die Funktion mit einer Taylorreihe in einem Punkt auf dem Hauptmeridian entwickelt, sind die Ableitungen nicht mehr von l abhängig.

Es ergeben sich Reihen der folgenden Form:

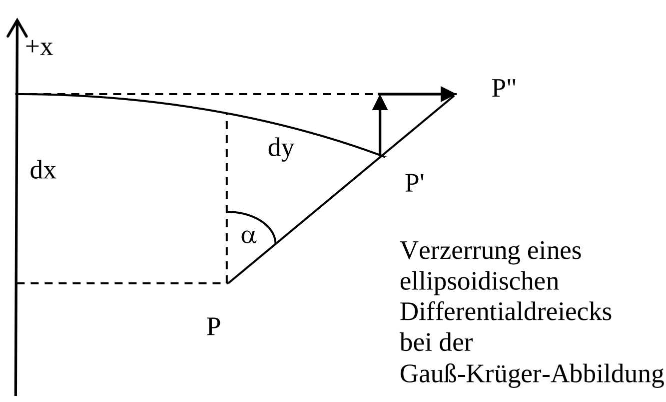
$$x+iy=F(q)+F'(q)(il)+\frac{1}{2}F''(q)(il)^2+\frac{1}{6}F'''(q)(il)^3+\cdots$$

Nach der Zerlegung der komplexen Reihe in Real- und Imaginärteil erhält man die gebräuchlichen Abbildungsgleichungen für die konforme Abbildung (Ableitungen siehe z.B. *Schnädelbach 1996, Großmann 1975*).



Durch die Abbildung des Ellipsoids in die Ebene ergeben sich Verzerrungen, die sich auf Längen, Flächen und Richtungen auswirken. Die Richtungsverzerrung ist bei der Gauß-Krüger-Abbildung wegen der Konformität so geringfügig, dass sie in der Praxis fast immer vernachlässigt werden kann.

Die Verzerrungen von Längen und Flächen können mittels Reduktionsformeln berechnet werden.

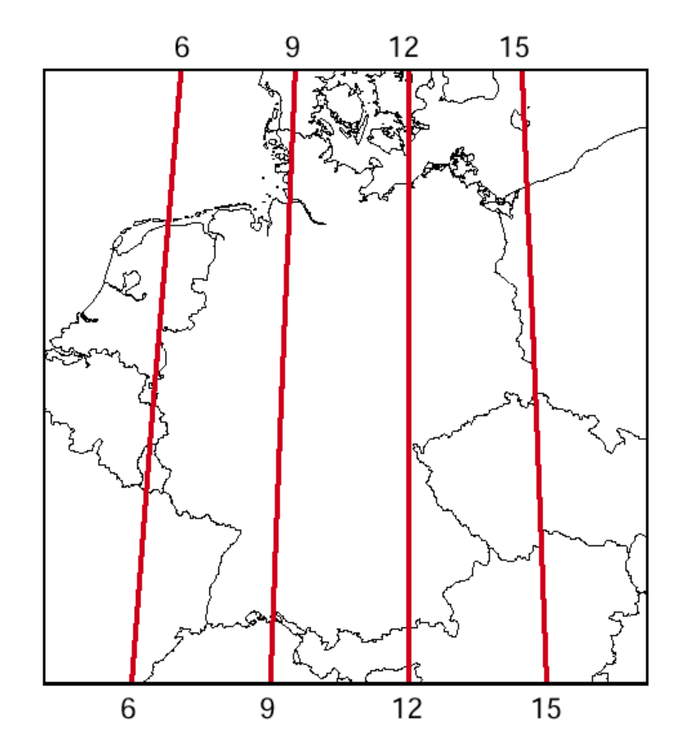


Verzerrung eines ellipsoidischen Differentialdreiecks bei der Gauß-Krüger-Abbildung

Die Längenreduktion zwischen gemessener Strecke S und aus Koordinaten gerechneter Strecke s beträgt näherungsweise: $\Delta S = s - S = \frac{y_m^2}{2R_m^2}S$

Die Flächenreduktion zwischen Feldfläche F und aus Koordinaten gerechneter Fläche f beträgt näherungsweise: $\Delta F = f - F = \frac{y_m^2}{R_m^2}F$

Amtliche Koordinaten in Bayern



Wegen der stark zunehmenden Längenverzerrung mit wachsendem Abstand (y) vom Hauptmeridian werden Gauß-Krüger-Meridianstreifensysteme von je 3° Längenunterschied gebildet.

In Bayern wird seit Einführung der Basis-GIS der Vermessungsverwaltung das gesamte Staatsgebiet einheitlich im 4. Meridianstreifen (Hauptmeridian: L_0=12°-Ost) abgebildet.

Der Abstand (y) vom Hauptmeridian beträgt in Bayern maximal etwa 220 km.

Bei y=200 km und R=6370 km ergibt sich ein Wert ΔS~0.5m für eine Strecke von 1km und ΔF~0.1ha für eine Fläche von $1km^2$ (100 ha).

Diese Beträge bei topographischen GIS-Anwendungen vernachlässigt werden.

Für behördliche GIS ist in Bayern entsprechend der gemeinsamen Bekanntmachung über den Aufbau raumbezogener Informationssysteme der Bayerischen Staatskanzlei und der Bayerischen Staatsministerien vom 7.1.1992 das amtliche Koordinatensystem als Raumbezugssystem zu verwenden.

Aus historischen und organisatorischen Gründen werden in Datenbeständen noch andere Bezugssysteme verwendet. Dies sind das bayerische Soldnersystem für alte Katasterdaten und die Flurkartenblattschnitte sowie das Universal Transversal Mercator System (UTM) für Daten aus dem militärischen Bereich oder dem Katastrophenschutz. Daten, deren Lageinformation Messungen mit dem Globalen Positionierungssystem (GPS) entstammt, die sich auf das World Geodetic System 1984 (WGS84) beziehen, gewinnen zunehmend an Bedeutung.

Die Arbeitgemeinschaft der Vermessungsverwaltungen in der Bundesrepublik (AdV) hat im Mai 1991 bei ihrer 88. Tagung beschlossen, dass zukünftig in der Bundesrepublik Deutschland für alle Aufgaben der Landesvermessung und des Liegenschaftskataster einheitlich das geodätische Bezugssystem ETRS89 (European Terrestrical Reference System) eingesetzt werden soll. Das ETRS89 ist eine genauere Definition des WGS84, für geodätische Zwecke, und stimmt für topographische Anwendungen mit dem WGS84 überein. Ergänzend wurde 1995 die Verwendung der UTM-Projektion für das neue amtliche Koordinatensystem festgelegt. Diese Festlegungen wurden 2007 im Rahmen der INSPIRE-Richtlinie für alle Mitgliedstaaten der EU übernommen.

Daraus folgt, dass die bisherigen amtlichen Gauß-Krüger-Koordinaten mit dem Datum Rauenberg (DHDN) zukünftig durch UTM-Koordinaten mit dem Datum ETRS89 ersetzt werden (AdV-Ziel: bis 201x). Diese Umstellung erstreckt sich erfahrungsgemäß über einige Jahrzehnte, da das gesamte Vermessungswerk davon betroffen ist, sie muss aber in der Konzeption einer langfristig angelegten Geodatenbank für die Bestandsdokumentation berücksichtigt werden. Bei einer Änderung des Raumbezugs müssen alle Koordinaten des Datenbestandes in das neue Bezugssystem transformiert werden. Zum Beispiel enthalten die Vektordatenbestände der DFK in Bayern ca. 100 Mio. Koordinaten zugrunde, das Fachgeoinformationssystems der Staatsforstverwaltung FORST-GIS enthält ca. 100 Mio. Koordinaten (12/2010), jährlich kommen etwa 5 Mio. neue Koordinaten dazu.

Hinweis auf Transformationen

Die Überführung verschiedener Koordinaten und Messwerte in das Bezugssystem eines GIS erfordert geeignete Transformationen, die einen hinreichend genauen Geokode (Raumbezug) liefern. Insbesondere soll die Aufnahmegenauigkeit durch Einflüsse von Transformationsfehlern nicht (wesentlich) verschlechtert werden. Die Genauigkeit der absoluten Lagerung der Datenbestände, d.h. ihres Raumbezuges, ist besonders für die Verschneidung verschiedener Themenbereiche wichtig, um signifikante Ergebnisse zu erzielen.

Innerhalb desselben Raumbezugssystems lassen sich je nach Erfordernis die Koordinaten in den verschiedenen Formen darstellen, wobei die **strenge Umformung** einer Darstellung in die andere möglich ist. So werden zum Beispiel in GIS als Geokode ebene Projektionskoordinaten bevorzugt, die Blattschnitte der Topographischen Karten sind mit geographischen Koordinaten festgelegt, für den räumlichen Übergang zwischen verschiedenen Bezugssystemen werden dreidimensionale kartesische Koordinaten benötigt.

Zwischen den Koordinaten eines Punktes in Raumbezugssystemen denen verschiedene gedoätische Bezugssysteme zugrunde liegen (z.B. GK und UTM) bestehen **keine strengen Beziehungen**, da den Koordinaten i.d.R. unterschiedliche Datumsfestlegungen und geodätischen Berechnungsgänge sowie unabhängige Beobachtungen zugrunde liegen. Die geometrischen Beziehungen zwischen den Koordinaten in Quell- und Zielsystem müssen daher empirisch ermittelt werden. Sie werden durch Transformationsgleichungen beschrieben, in denen die gesuchten Koordinaten des Zielsystems als Funktion der gegebenen Koordinaten des Quellsystems mit den zu bestimmenden Parametern dargestellt werden. Das Modell der Transformation muss im Transformationsgebiet die geometrischen Beziehungen zwischen den beiden Systemen möglichst gut repräsentieren. Solche Transformationen bewirken einen translativen oder projektiven **Datumsübergang**, da primär der Wechsel des Bezugssystems erfolgt. Einflüsse der unterschiedlichen Datumsfestlegungen und von Netzverzerrungen können dabei jedoch nicht getrennt werden.

Für Datumsübergänge ist daher zu empfehlen, bekannte Einflüsse von Netzverzerrungen, z.B. durch die Projektion, vor Ermittlung der Transformationsparameter zu beseitigen. Netzverzerrungen aufgrund unterschiedlicher Messgenauigkeit und Messfehlern zeigen sich als Restklaffungen (Widersprüche) in den Passpunkten.

Restklaffungen können bei Bedarf mit Interpolationsansätzen weiterverarbeitet werden, man spricht dann von der Homogenisierung der Koordinaten.

Folgende Koordinatenumformungen und Transformationen sind in GIS von Bedeutung →



Koordinatenumformung
3D-Kartesische P(X,Y,Z) in Geodätische Koordinaten (B,L,h)
Geodätische (B,L) in ebene konforme Projektionskoordinaten (y,x)
Konforme Projektionskoordinaten (y,x) in Gebrauchskoordinaten (R,H)
Datumstransformation
Datums-Shift nach Helmert, für kleine Drehungen und Maßstab nahe bei 1 (6 Parameter, linear konform) 3D-Datumsübergang
Höhenübergang Ellipsoid-/NN-Höhen 1D-Datumsübergang
Ähnlichkeitstransformation (Helmerttransformation) (4 Parameter, linear konform) 2D-Datumsübergang
Affintransformation (6 Parameter) 2D-Datumsübergang

7 Selektion, Analyse und Präsentation *(Überblick)*

Nach *Bill/Fritsch (1991)* umfasst ein GIS Funktionen *(EVAP)* für die Erfassung, Verwaltung, Analyse und Präsentation von Geodaten. In den vorangegangen Abschnitten wurden die Modellierung, die Erfassung und Verwaltung von Geodaten in GIS besprochen.

Zur (Wieder-)Gewinnung der Geoinformation ist es jedoch erforderlich die Geodaten in Abhängigkeit von der jeweiligen Fragestellung zu:

- selektieren (auswählen) → Teilmenge bilden
- analysieren (verarbeiten) → Geoprocessing
- präsentieren (darstellen) → Report erstellen

Für diese Aufgaben sind im GIS Operationen und Funktionen erforderlich, mit denen Geodaten in allen grundlegenden Datenstrukturen verarbeitet und in Beziehung gesetzt werden können. Die grundlegenden Datenstrukturen für Geodaten sind:

- Attribute (Sachdaten)
- Geometrieelemente (Vektoren, Features)
- Rasterdaten (Bild, Grid)

In der Leistungsfähigkeit und Vielfalt der Werkzeuge (Tools) für die Analyse und Präsentation von Geodaten unterscheiden sich die gängigen GIS-Softwareprodukte wesentlich. Der Umfang und die Mächtigkeit der verfügbaren Tools ist neben den Möglichkeiten zur Datenmodellierung ein primäres Leistungsmerkmal eines GIS-Softwareproduktes, das auch wesentlich seinen Preis bestimmt. So bietet ESRI bei seinem Standardprodukt ArcGIS-Desktop die Leistungsstufen: ArcView, ArcEditor und ArcInfo, die sich in der zunehmenden Mächtigkeit ihrer Toolbox (und im Preis) unterscheiden.

Nachfolgend wird eine Auswahl von grundlegenden Operationen und Funktionen für die Verarbeitung von Geodaten „core tools“ vorgestellt, die repräsentativ sind für viele GIS-Softwareprodukte. In den Übungen werden diese an praktischen Beispielen mit ArcGIS angewendet und ausführlich erläutert. Hier werden sie, als Ergänzung zu den Übungen, kurz zusammengefasst. In den Übungen wird die Leistungsstufe ArcView eingesetzt, die alle „core tools“ breitstellt.

Für die Erläuterung der Operationen und Funktionen wird auf den Übungsstoff verwiesen und auf die ArcGIS-Hilfe zurückgegriffen. Eine umfassende Übersicht aller verfügbaren Werkzeuge findet man auch in der ESRI-Produktdokumentation für ArcGIS Desktop.

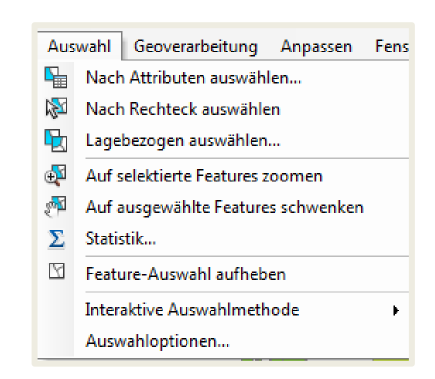
(siehe: Übungen zu den Modulen GI und GIS sowie z.B. Geoprocessing Commands Quick Reference Guide, ArcGIS Desktop core geoprocessing tools)



Selektion, die Auswahl von Geodaten

In GIS-Software kann man drei grundlegende Selektionsmethoden unterscheiden, deren Ergebnismengen man mittels logischer Verknüpfungen kombinieren kann:

- Auswahl nach Attributen (Sachdaten)
- Auswahl durch die Eingabe von Suchgeometrien (Geometrie)
- Auswahl über räumliche Relationen, lagebezogen (Topologie)



Die Auswahl von Geodaten mittels Prädikaten über die Attributwerte sowie Suchgeometrien und räumliche Relationen ist in GIS-Produkten unterschiedlich realisiert. Für die verschiedenen Selektionsmethoden gibt es meist eigene Werkzeuge, die eine schrittweise Auswahl und logische Verknüpfung der Teilauswahlen ermöglichen.

Während der Sachdatenzugriff *(„Nach Attributen auswählen“)* meist mittels SQL-Select-Anweisungen erfolgt, sind die Geometriezugriffe herstellerspezifisch (propritär) realisiert. Für die geometrische Auswahl von Geodaten sind die Punktidentifikation oder die Auswahl in Rechtecks- oder Polygonbereichen üblich. Zunehmend erlauben immer mehr Produkte auch die Selektion mittels topologischer Relationen, d. h. über räumliche Beziehungen zwischen Geoobjekten. Bei der geometrisch/topologischen Selektion erfolgt die Auswahl über die primäre Eigenschaft der Geodaten, ihren Raumbezug.

Geographic Query Language: Im Rahmen der Normierung (ISO) und Standardisierung (OGC) von GIS-Produkten, wurde auch eine Abfragesprache entworfen, die es ermöglichen soll in Datenbanken Abfragen nach Attributen mit geometrischen und topologischen Bedingungen zu kombinieren. Diese erweiterte Abfragesprache (SQL/MM-Spatial, ISO 13249) wird auch als GQL oder SQL-Spatial bezeichnet (OGC, Praxisanforderung). Sie ist eine Erweiterung des Standard-SQL (SQL 3, ISO 9075) um räumliche Funktionen und Operatoren. SQL-Spatial steht heute in einigen Datenbankprodukten zur Verfügung (z.B. Oracle, PostGIS), ist aber sehr unterschiedlich realisiert.

Auswahl nach Attributen (Sachdaten)

Die Auswahl über Attribute, die die semantischen Merkmale von Geoobjekten beschreiben, erfolgt meist mittels Select-Anweisungen, in denen Aussagen (Prädikate) über die Attributwerte gesuchter Geoobjekte gemacht werden.

Aufbau einer Select-Anweisung (SQL)

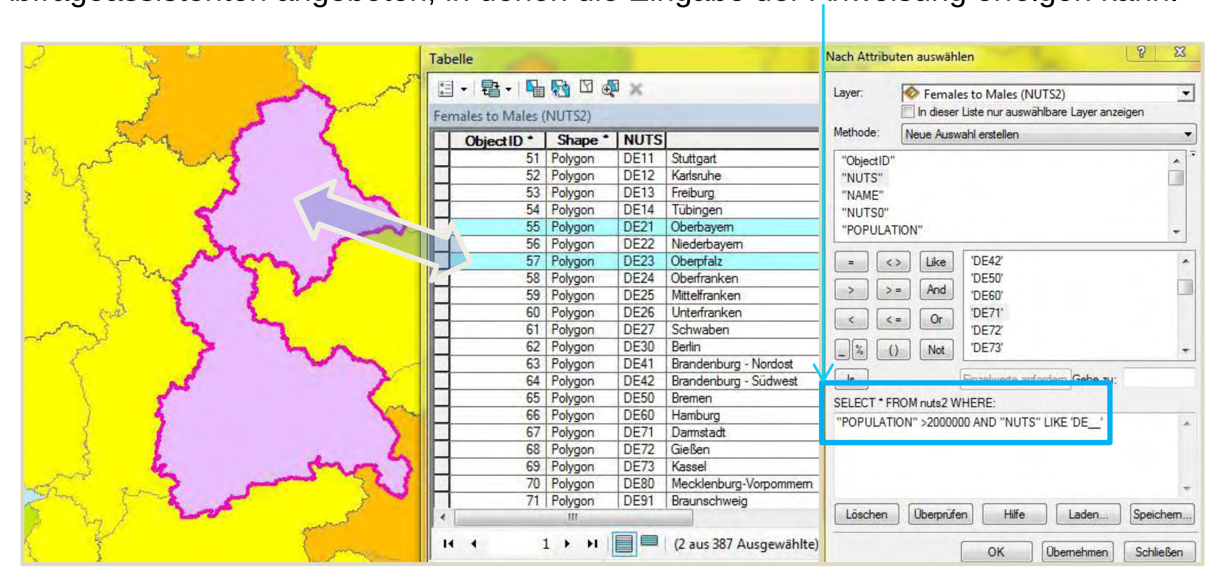
<table>
<tr><td>SELECT {ergebnisspalten}</td><td>SPALTENAUSWAHL
(Prokjektion, *=alles)</td><td rowspan="5">WHERE-Bedingung
Prädikate über Attributwerte
prädikat : ausdruck {op} ausdruck
ausdruck {op} select ...
op : (=,<,>,<=,>=,<>)
(in, not in)
(like, not like)
logisch verknüpfte prädikate :
prädikat(1) {and|or} prädikat(2)
ausdruck : attributwert, liste,
funktion, konstante</td></tr>
<tr><td>FROM tabelle [tabsynonym], ...</td><td>TABELLENAUSWAHL
(Tabellenverbund)</td></tr>
<tr><td>[WHERE prädikate]</td><td>ZEILENAUSWAHL
(Selektion)</td></tr>
<tr><td>[GROUP BY attribut, ...]
[HAVING bedingung]</td><td>GRUPPIERUNG
(Aggregate)</td></tr>
<tr><td>[ORDER BY spalte, ...]</td><td>SORTIERUNG</td></tr>
</table>

Geoobjekte entsprechen einer eindeutig identifizierbaren Entität des GeoRaumes, es sind Informationseinheiten, die aus Geometrie- und Sachdaten gebildet werden. Damit dies möglich ist muss zwischen diesen beiden Datentypen zur Modellierung von Geoobjekten, den Sach- und Geometriedaten, eine Verbindung bestehen.

Im Shape-Format werden die Attribute (Sachdaten) eines Themas in einer Tabelle vorgehalten, die über eine spezielle, vom System gepflegte Relation, mit den Geometriedaten (Punkt- | Linien - | Flächen-Feature) verknüpft sind.

Shape-Thema = [Geometie (shp) ⇔ Geolink (shx) ⇔ Tabelle (dbf)]

Für die Auswahl von Geodaten nach Attributen werden meist Select-Anweisungen ohne Projektion verwendet, die nur eine *„where"*-Klausel mit den Prädikaten über die Attributwerte der gesuchten Geoobjekte enthalten. Dazu werden von den GIS-Produkten Abfrageassistenten angeboten, in denen die Eingabe der Anweisung erfolgen kann.



Beispiel Abfrageassistent von ArcGIS „Nach Attributen auswählen"
*Select * from <Thema>.dbf* ***where*** *<Pädikate>*
Bei der Abfrage werden immer Geoobjekte ausgewählt, mit ihren Geometrie- und Sachdaten.
Die ausgewählten Geoobjekte werden in Tabelle und Geometrie (Thema) markiert.

Auswahl durch markieren von Zeilen direkt in der Tabelle

Neben der Selektion mit einer Select-Anweisung kann in speziellen Fällen auch die direkte Auswahl der Geoobjekte, durch die Markierung der Datensätze in der Tabelle sinnvoll sein. Dazu kann man einen Datensatz mit dem Cursor in der Auswahlspalte (graue Felder) der Tabelle markieren. Durch die folgenden Tastenkombinationen kann die Auswahl auch erweitert bzw. reduziert werden.

Auswahl durch Markierten einer Zeile mit der Cursorposition

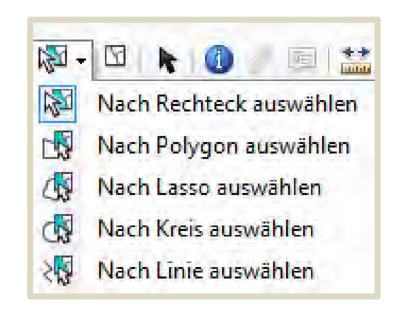
Auswahlbereich durch das Festhalten der Shift-Taste angeben, eine bestehende Auswahl erweitern

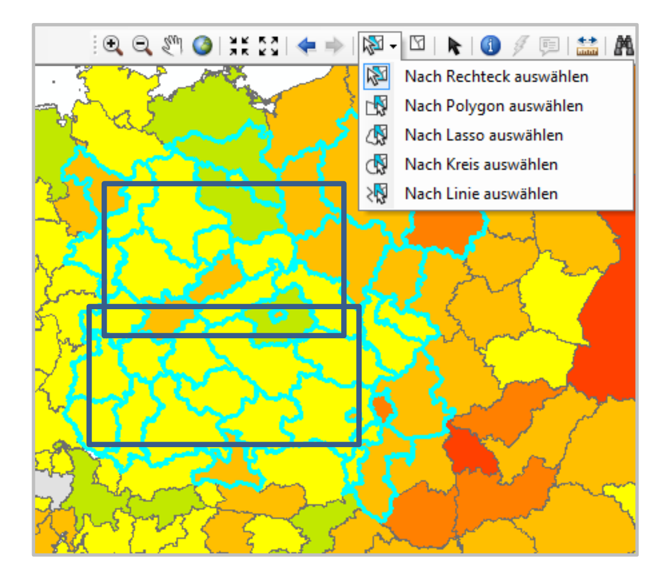
Auswahl wird durch das Festhalten der Strg-Taste um markierte Sätze reduzieren Strg

In ArcGIS werden Funktionen und Operationen nur auf die ausgewählten (markierten) Geoobjekte eines Themas (feature class) angewendet. Falls alle Geoobjekte bearbeitet werden sollen, müssen alle ausgewählt sein oder keines. Die Auswahlmenge muss ggf. vor der Durchführung einer Auswertung mit der Funktion *„Feature-Auswahl aufheben"* zurückgesetzt werden, wenn alle Geoobjekte bearbeitet werden sollen.

Auswahl mit Geometrieeingaben (Positionen, Bereiche)

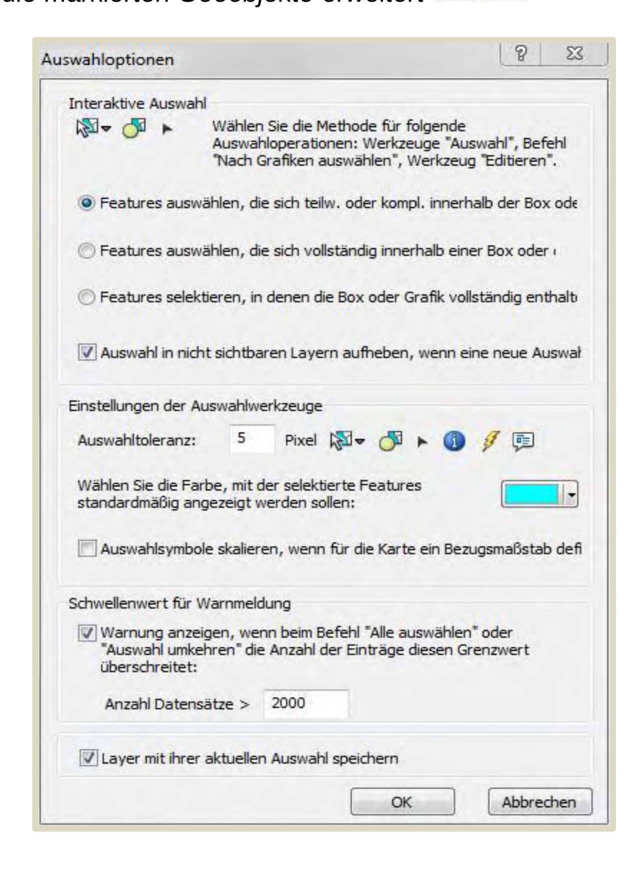
Die Auswahl von Geoobjekten über eine Geometrie erfolgt durch die Eingabe von Positionen, Rechtecken oder Polygone mit dem Cursor im View-Fenster, das die grafische Repräsentation der Geoobjekte aus den Themen enthält.
Die Selektion *„Nach <Geometrie> auswählen“* geht über alle dargestellten / ausgewählten Layer in einem Datenrahmen.





Auswahl durch markieren der Geoobjekte mit der Cursorposition

Die Auswahl wird beim Drücken der Shift-Taste um die markierten Geoobjekte erweitert



Die wichtigste Variante für die geometrische Auswahl ist die Funktion *„Nach Rechteck auswählen“*, bei der achsparallele Rechtecke (Box) oder Einzelpositionen mit dem Cursor eingegeben werden können.

Die geometrische Auswahl kann mit Hilfe von Optionen gesteuert werden. Die wichtigsten sind die Interpretation der Box und die Auswahltoleranz für Positionen.

Für die Sichtbarkeit der Auswahl im View, sollte eine gut erkennbare Farbe für die Markierung verwendet werden.

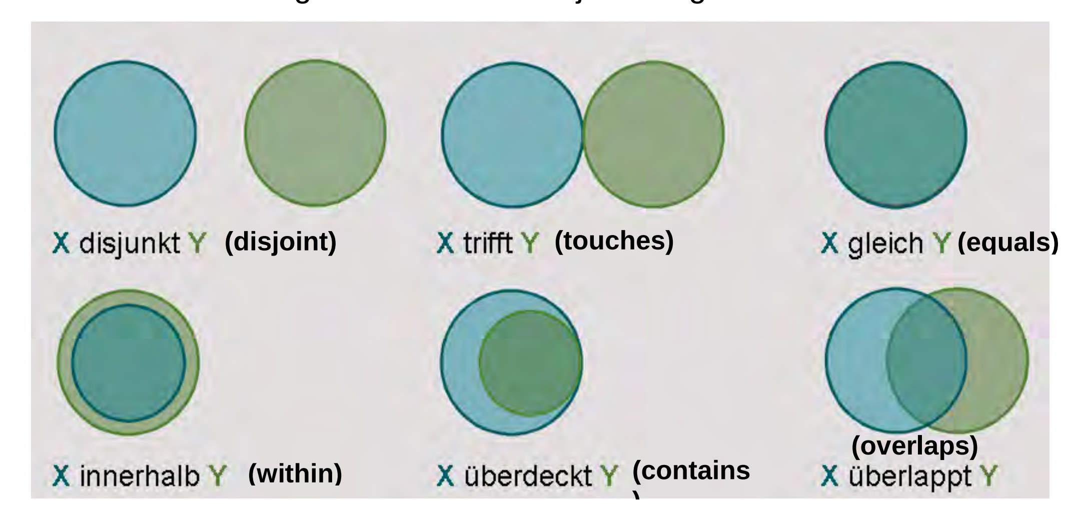
Auswahl über räumliche Relationen (Topologie)

Die Auswahl von Geoobjekten über räumliche Relationen *„Lagebezogen auswählen ….“* ist die wichtige Möglichkeit in GIS den Raumbezug und die daraus resultierenden räumlichen Relationen zwischen verschiedenen Geoobjekten direkt zu nutzen. Man kann mit dieser Selektionsmöglichkeit Aufgaben lösen, bei denen die Nähe, Umgebung, und Umschließung der Geoobjekte eine Rolle spielen.

Grundlage sind die topologischen Invarianten und die daraus ableitbaren Relationen. Toplogischen Invarianten im zweidimensionalen Raum sind die Geschlossenheit, Schnittpunktreue, Trennung innen/außen und die Randpunkteigenschaft (siehe 4-16). Es werden Operatoren benötigt, die die Geometrie zweier Gruppen von Geoobjekten bezüglich eines gewählten Kriteriums vergleichen und als Ergebnis TRUE zurück liefern, wenn diese Relation erfüllt ist, sonst FALSE.

Dimensionally 4-Intersection Model (D-4IM)

Als Grundlage für die Definition topologischer Operatoren sollen zwei Flächenobjekte X und Y auf Basis des Dimensionally 4-Intersection Model (D-4IM) betrachtet werden, für Geoobjekte mit gleichem Geometrietyp (Flächen). Es lassen sich 6 grundlegende Fälle für die räumliche Beziehung der beiden Geoobjekte angeben.



Für die Geoobjekte X und Y sollen ihr Rand (δX, δY) und ihr Inneres (^{O}X,^{O}Y) getrennt betrachtet werden, so ergeben sich aus den 4 Kombinationen (TRUE oder FALSE) Muster (Variation) die den gesuchten Operatoren entsprechen.

Operator	$\delta X \cap \delta Y$	$\delta X \cap {}^{O}Y$	$^{O}X \cap \delta Y$	$^{O}X \cap {}^{O}Y$
disjoint	F	F	F	F
touches (meets)	T	F	F	F
equals	T	F	F	T
within (inside)	F	T	F	T
contains	F	F	T	T
overlaps	T	T	T	T

Dimensionally Extended 9-Intersection Model (DE-9IM)

Die Verallgemeinerung des D-4IM für Operatoren, die Geometrieelemente unterschiedlicher Dimension verarbeiten (Punkt|Linie|Fläche), ergibt das auf Egenhofer zurückgehende Dimensionally Extended 9-Intersection Model (DE-9IM), das auch als Neuner-Verschneidungs-Matrix (NVM) bezeichnet wird.

Das DE-9IM (NVM) nach Egenhofer beschreibt die topologischen Beziehungen zwischen zwei Geometrieelementen.

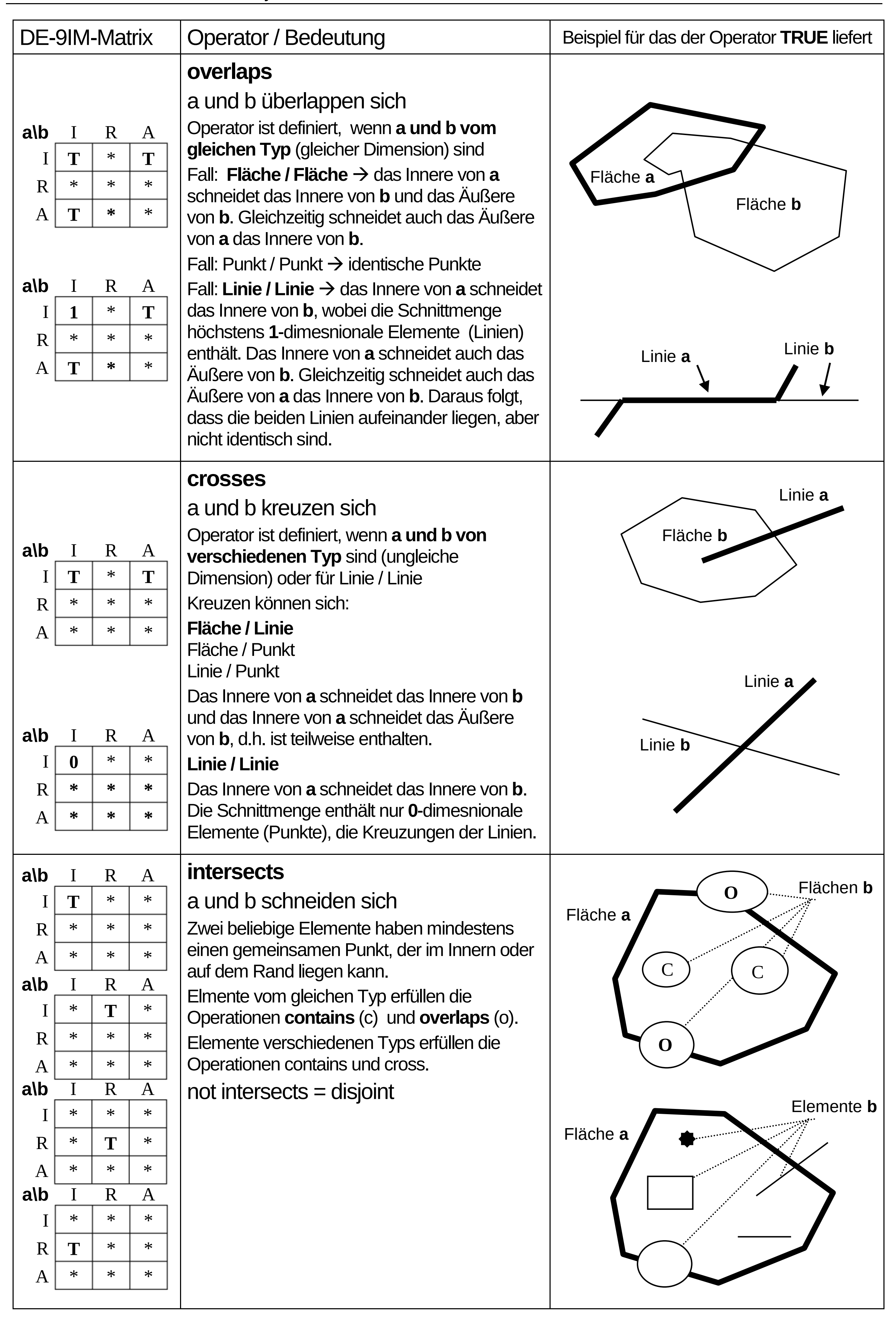
Jede Geometrie unterteilt den 2-dimensionalen Raum in drei Bereiche, das Innere (interior), den Rand (boundary) und das Äußere (exterior).

DE-9IM X_{ab}		Element **b**		
		Inneres	Rand	Äußeres
Element **a**	Inneres	X_{II}	X_{IR}	X_{IA}
	Rand	X_{RI}	X_{RR}	X_{RA}
	Äußeres	X_{AI}	X_{AR}	X_{AA}

Jede topologische Beziehung zwischen zwei Geoobjekten kann als das Ergebnis der neun möglichen Schnittmengen ihrer Geometrie, die zwischen dem Inneren (I), dem Rand (R) und dem Äußeren (A) der beiden Objekte unterschieden wird, beschrieben werden.

Übersicht: Egenhofer-Operatoren (DE-9IM)

DE-9IM-Matrix	Operator / Bedeutung	Beispiel für das der Operator **TRUE** liefert
a\b I R A I **T** * * R * * * A **F** **F** *	**contains** a enthält b vollständig Das Innere von **a** schneidet das Innere von **b** und das Äußere von **a** schneidet weder das Innere noch den Rand von **b**.	Fläche **a** Fläche **b**
a\b I R A I **T** * **F** R * * **F** A * * *	**within (inside)** a ist vollständig innerhalb von b Das Innere von **a** schneidet das Innere von **b** und das Äußere von **b** schneidet weder das Innere noch den Rand von **a**. a contains b = b within a	Fläche **b** Fläche **a**
a\b I R A I **F** **F** * R **F** **F** * A * * *	**disjoint** a und b berühren sich nicht Die Schnittmengen zwischen dem Inneren und dem Rand von **a** und **b** sind leer, es gibt keine gemeinsame Punkte. not disjoint = intersects	Fläche **a** Fläche **b**
a\b I R A I **T** * **F** R * * **F** A **F** **F** *	**equals** a und b sind (geometrisch) gleich Das Innere von **a** und **b** schneidet sich. Das Äußere von **a** schneidet weder das Innere noch den Rand von **b.** Das Äußere von **b** schneidet weder das Innere noch den Rand von **a**. **a** und **b** sind kongruent (innerhalb Toleranz)	Fläche **a** = Fläche **b**
a\b I R A I **F** **T** * R * * * A * * * OR **a\b** I R A I **F** * * R **T** * * A * * * OR **a\b** I R A I **F** * * R * **T** * A * * *	**touches (meets)** a und b berühren sich am Rand Die Schnittmenge zwischen dem Inneren von **a** und **b** ist leer. Jedoch schneidet entweder der Rand von **b** das Innere von **a** oder der Rand von a das Innere von **b** oder die Ränder von **a** und b schneiden sich. Berühren können sich: Fläche / Fläche Fläche / Linie Fläche / Punkt Linie / Linie Linie / Punkt Nicht für Punkt / Punkt definiert, da Punkte keinen Rand haben.	Fläche **b** Punkt **a** Linie **b** Linie **a** Fläche **a** Fläche **b**



DE-9IM-Matrix	Operator / Bedeutung	Beispiel für das der Operator **TRUE** liefert

overlaps

a und b überlappen sich

a\b	I	R	A
I	**T**	*	**T**
R	*	*	*
A	**T**	*	*

a\b	I	R	A
I	**1**	*	**T**
R	*	*	*
A	**T**	*	*

Operator ist definiert, wenn **a und b vom gleichen Typ** (gleicher Dimension) sind

Fall: **Fläche / Fläche** → das Innere von **a** schneidet das Innere von **b** und das Äußere von **b**. Gleichzeitig schneidet auch das Äußere von **a** das Innere von **b**.

Fall: Punkt / Punkt → identische Punkte

Fall: **Linie / Linie** → das Innere von **a** schneidet das Innere von **b**, wobei die Schnittmenge höchstens **1**-dimesnionale Elemente (Linien) enthält. Das Innere von **a** schneidet auch das Äußere von **b**. Gleichzeitig schneidet auch das Äußere von **a** das Innere von **b**. Daraus folgt, dass die beiden Linien aufeinander liegen, aber nicht identisch sind.

crosses

a und b kreuzen sich

a\b	I	R	A
I	**T**	*	**T**
R	*	*	*
A	*	*	*

a\b	I	R	A
I	**0**	*	*
R	*	*	*
A	*	*	*

Operator ist definiert, wenn **a und b von verschiedenen Typ** sind (ungleiche Dimension) oder für Linie / Linie

Kreuzen können sich:

Fläche / Linie
Fläche / Punkt
Linie / Punkt

Das Innere von **a** schneidet das Innere von **b** und das Innere von **a** schneidet das Äußere von **b**, d.h. ist teilweise enthalten.

Linie / Linie

Das Innere von **a** schneidet das Innere von **b**. Die Schnittmenge enthält nur **0**-dimesnionale Elemente (Punkte), die Kreuzungen der Linien.

intersects

a und b schneiden sich

a\b	I	R	A
I	**T**	*	*
R	*	*	*
A	*	*	*

a\b	I	R	A
I	*	**T**	*
R	*	*	*
A	*	*	*

a\b	I	R	A
I	*	*	*
R	*	**T**	*
A	*	*	*

a\b	I	R	A
I	*	*	*
R	**T**	*	*
A	*	*	*

Zwei beliebige Elemente haben mindestens einen gemeinsamen Punkt, der im Innern oder auf dem Rand liegen kann.

Elmente vom gleichen Typ erfüllen die Operationen **contains** (c) und **overlaps** (o).

Elemente verschiedenen Typs erfüllen die Operationen contains und cross.

not intersects = disjoint

Die Einträge X_{ab} der Matrix können die Werte { T | F | * | 0 | 1 } annehmen. Ein Operator ist genau dann erfüllt, wenn alle Teilbedingungen X_{ab} seiner entsprechenden Matrix erfüllt sind.

Symbol	Bedeutung	Erklärung der möglichen Wete für X_{ab}
T	Wahr (true)	**Schnittmenge ist nicht leer**, aber von beliebiger Dimension → kann aus Punkten, Linien oder Flächen bestehen
F	Falsch (false)	**Schnittmenge ist leer**
*****	beliebig	**Schnittmenge nicht von Bedeutung** für das Ergebnis eines Operators
0	O-dimensional	**Schnittmenge** mit Objekten, deren maximale **Dimension 0** ist → sie enthält einen oder mehrere Punkte
1	1-dimensional	**Schnittmenge** mit Objekten, deren maximale **Dimension 1 i**st → sie enthält eine oder mehrere Linien, auch Punkte sind erlaubt

Grundlegende Topologische Operatoren nach dem DE-9IM:

contains, **within** (inside), **overlaps**, **crosses**, **intersects**, **disjoint**, **equals**, **touches** (meets)

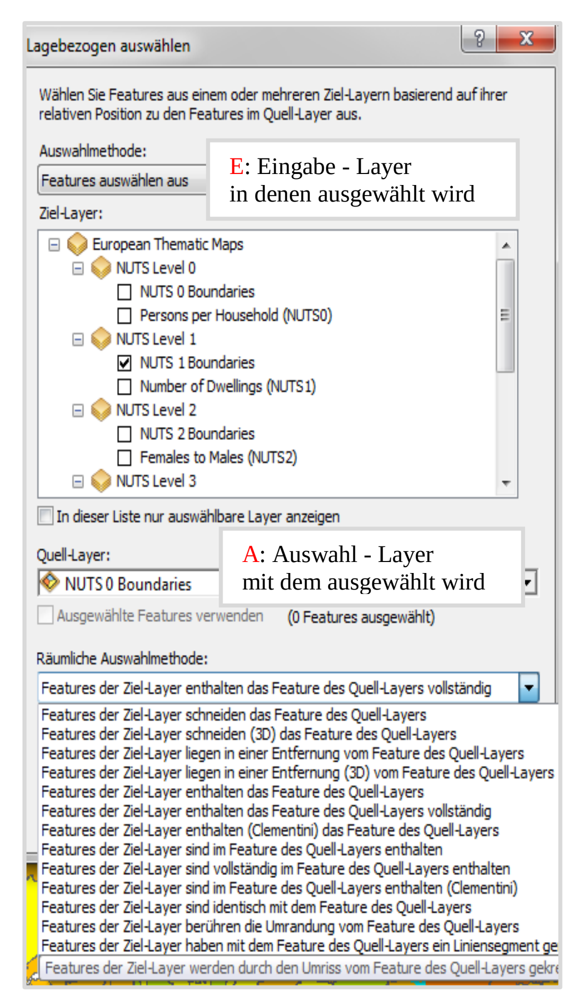
In einigen Implementierungen gibt es zusätzlich den allgemeinen Operator „**relate**", dem direkt eine (3*3)-Matrix, in Form einer neunstelligen Bitfolge übergeben werden kann. Damit kann der Anwender beliebige Relationen auswerten. Zusätzliche Operatoren ergeben sich, wenn man ausgedehnte Objekte (Flächen) durch ihren Bezugspunkt repräsentiert.

Lagebezogen auswählen in ArcGIS,

in Anlehnung an ArcGIS-Hilfe (Eingabe-Layer in dem ausgewählt wird | Auswahl-Layer mit dem ausgewählt wird):

Für die Auswahlfunktion "Lagebezogen auswählen" sind in ArcGIS folgende räumliche Operatoren verfügbar (Auswahl):

- **sich überschneiden mit** (*A intersects E*) – Die Features in den Eingabe-Layern werden ausgewählt, wenn sie sich mit einem Feature im Auswahl-Layer überschneiden.
- **vollständig enthalten** (*A within E*) – Die Features in den Eingabe-Layern werden ausgewählt, wenn sie ein Feature im Auswahl-Layer vollständig enthalten. Bei dem Auswahl-Layer muss es sich um einen Polygon-Layer handeln.
- **enthalten** – Die Features in den Eingabe-Layern werden ausgewählt, wenn sie ein Feature im Auswahl-Layer enthalten.
- **vollständig enthalten sind in** (*A contains E*) – Die Features in den Eingabe-Layern werden ausgewählt, wenn sie sich vollständig innerhalb eines Features im Auswahl-Layer befinden bzw. darin enthalten sind. Der Auswahl-Layer muss ein Polygon-Layer sein, oder es muss ein Puffer rund um die Punkt- und Linien-Features vorhanden sein.
- **enthalten sind in** – Die Features in den Eingabe-Layern werden ausgewählt, wenn sie in einem Feature im Auswahl-Layer enthalten sind. Bei dem Auswahl-Layer muss es sich um einen Polygon-Layer handeln.
- **identisch sind zu** (*A equals E*) – Die Features in den Eingabe-Layern werden ausgewählt, wenn ihre geometrische Form mit der eines Features im Auswahl-Layer übereinstimmt.
- die Umrandung berühren von (*A touches E*) – Die Features in den Eingabe-Layern werden ausgewählt, wenn ihre Grenze ein Feature im Auswahl-Layer berührt. Bei Eingabe- und Auswahl-Layern muss es sich um Linien oder Polygone handeln.



- **ein Liniensegment gemeinsam haben mit** – Die Features in den Eingabe-Layern werden ausgewählt, wenn sie mit einem Feature im Auswahl-Layer ein gemeinsames Liniensegment teilen. Bei den Eingabe-Layern und dem Auswahl-Layer muss es sich um Linien oder Polygone handeln.
- **gekreuzt werden durch den Umriss von** (*A crosses E*) – Die Features in den Eingabe-Layern werden ausgewählt, wenn sie vom Umriss eines Features im Auswahl-Layer gekreuzt werden. Bei den Eingabe-Layern und dem Auswahl-Layer muss es sich um Linien oder Polygone handeln.
- **ihren Mittelpunkt haben** in – Die Features in den Eingabe-Layern werden ausgewählt, wenn ihr Bezugspunkt innerhalb eines Features im Auswahl-Layer liegt.
- **in einer Entfernung liegen von** – Die Features in den Eingabe-Layern werden ausgewählt, wenn sie sich in einem bestimmten Abstand zu einem Feature im Auswahl-Layer befinden. Bei Auswahl dieser Option wird das Feld "Pufferentfernung" am unteren Rand des Dialogfelds automatisch aktiviert, sodass Sie die Entfernung angeben können. Über die Option Pufferentfernung in diesem Dialogfeld kann man die Entfernung angeben, die für den Operator in einer Entfernung liegen von verwendet wird.

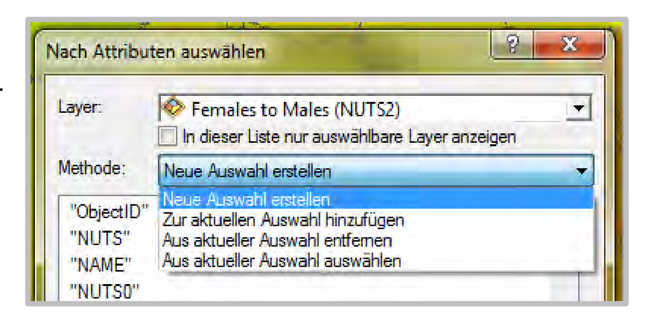
Bei vielen Operatoren, insbesondere equals und touches, sind wegen der Unschärfe der Metrik Puffertoleranzen zu verwenden. Die Features im Auswahl-Layer erhalten dann einen Puffer mit einer bestimmten Breite. Wenn die Auswahl vorgenommen wird, werden dann statt der Featurekonturen die Pufferregionen um die Features im Auswahl-Layer verwendet.

Kombination von Teilabfragen

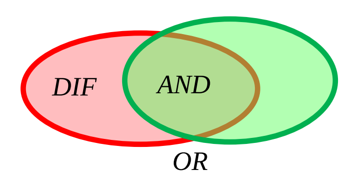
Für Selektionen in GIS sind häufig attributive und räumliche Abfragen in Kombination erforderlich, damit die Auswahl von Geodaten nach allen relevanten Merkmalen erfolgen kann. Die Formulierung solcher Abfragen in geschlossener Form, wie mit GQL in Datenbanken, ist in GIS-Softwareprodukten nicht realisiert. Für die kombinierte Abfrage in ArcGIS sind Teilabfragen zu bilden. Die Ergebnisse dieser Teilabfragen werden durch Markierung der ausgewählten Geoobjekte gemerkt. Mehre Teilabfragen können dann über logische Operatoren verknüpft werden, die die vorhandene (gemerkte) Auswahl mit der aktuellen Auswahl verknüpfen.

In ArcGIS stehen dazu in den Assistenten für „Auswahl nach Attributen“ und „Lagebezogene Auswahl“ folgende Möglichkeiten zur Verknüpfung von Teilauswahlergebnissen zur Verfügung:

- Zur aktuellen Auswahl hinzufügen (**OR**-Verknüpfung → verallgemeinert)
- Aus der aktuellen Auswahl auswählen (**AND**-Verknüpfung → spezialisiert)
- Aus der aktuellen Auswahl entfernen



aktuelle Auswahlmege A, gemerkte (vorhandene) Auswahlmenge M

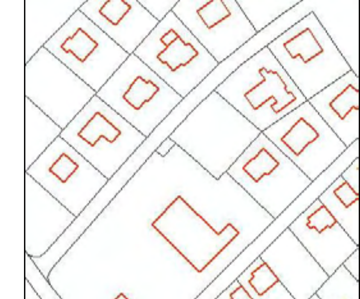


$OR: \; A \cup M$ A vereinigt mit M

$AND: A \cap M$ A geschnitten mit M

$DIF: \; A \setminus M$ A minus M

Die Möglichkeit zu Zerlegung komplexer Abfragen nach Geometrie/Topologie und Attributen und der anschließenden logischer Verknüpfung der Teilabfragen ist aus Anwendersicht meist ausreichend und i.d.R. auch einfacher zu formulieren als eine geschlossene Abfrage mit GQL.



Beispiel: *in den Themen Haus und Flurstück → alle bebauten Flurstücke in der ‘Karlstraße‘, die größer als 1000 m² sind suchen.*

*select * from Haus.dbf where straße=‘Karlstraße‘ (räumlich einschräken)*
Haus.shp within Flurstück.shp (nur ausgewählte Feature verwenden)
*select * from Flurstück.dbf where fläche>1000 (aus aktueller Auswahl auswählen)*

Geoprocessing (Analyse)

Man kann die folgenden grundlegenden Analysefunktionen eines GIS unterscheiden:

- Messen, Zählen und Berechnen (z.B. Flächen, Längen) sowie deskriptive statistische Verfahren (z.B. Mittelwerte, Standardabweichungen, Min/Max-Werte)
- Verschneidung und Aggregation von Themen
- Räumliche Beziehungsanalysen, u.a. Zonen- und Puffergenerierung, Interpolation
- Räumliche Statistik
- Simulationen, Szenarienberechnung

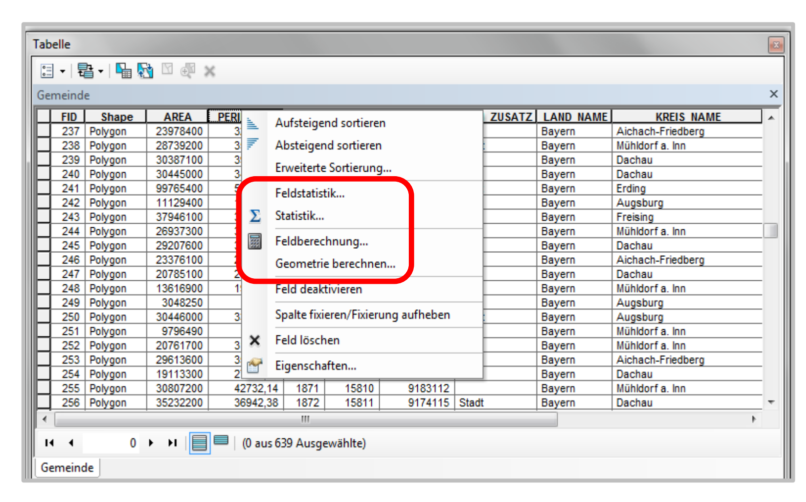
Am Beispiel der Geoprocessing-Tools von ArcGIS werden nachfolgend einige grundlegende Operationen und Funktionen aufgezeigt. Alle Auswertungen in ArcGIS beziehen sich immer auf die ausgewählten Geoobjekte.

Funktionen zur Berechnung neuer Attribute

Funktionen bestimmen Eigenschaften von Geoobjekten und liefern als Ergebnisse Attributwerte, z. B. die Längen von Linien oder die Größen von Flächen.

Berechnung von Attributwerten

Die Berechnung von Attributwerten lässt sich in „Feldberechnung“ und „Geometrie berechnen“ unterscheiden, je nachdem ob neue Werte aus den Attributwerten (dbf) oder den metrischen Merkmalen (shp) der Geoobjekte berechnet werden.



Statistik für Attribute

Für numerische Attribute ist es hilfreich, wenn man Aussagen über die Verteilung ihrer Werte und den Wertebereich machen kann.

Hierzu gibt es in ArcGIS die Funktion „Statistik“.

Einen guten, grafischen Überblick zu Attributwerten kann man sich auch mit Hilfe der Symologie schnell verschaffen (siehe Präsentationsfunktionen).

Feldstatistik für Attribute

Für die Aggregation (Summe, Mittelwert, Minimum, Maximum, ...) von Attributwerten steht die Funktion „**Feldstatistik**“ zur Verfügung. Diese Funktion entspricht der *„group by“*-Klausel in einer SQL-Select-Anweisung. Aggregatfunktionen liefern für eine Gruppe genau einen Wert ab, zu einer Gruppe werden alle Sätze zusammengefasst, die gleiche Werte im Gruppierungsattribut (Aggregationsattribut) haben.

Beispiel für eine Feldstatistik mit SQL:

SELECT	**Landkreis**, sum(Area), avg(Area)	Projektion mit Aggregaten
FROM	Gemeinde.dbf	
[WHERE	Land_Name = ‘Bayern‘]	Selektion zur Zeilenauswahl
GROUP BY	**Landkreis**	Aggregation (Feldstatistik)

(vorher ggf. Zeilenauswahl mit „Nach Attributen auswählen“, dann die Gruppierung mit der Funktion „Feldstatistik“ durchführen)

Die Anwendung der Feldstatistik ergibt eine Gruppierung nach dem ausgewählten Attribut (Feld) und liefert für diese Gruppen die Werte der angegebenen Aggregate in einer neuen Ergebnistabelle, die dann auch in Form von Diagrammen grafisch dargestellt werden können.

(Einzelheiten zur Anwendung dieser Funktionen in ArcGIS werden in den Übungen behandelt)

Geometrische Operatoren zur Auswertung von Geoobjekten (ArcView)

Es wird eine Auswahl wichtiger Operatoren für die Verarbeitung von Geodaten angegeben, die grundlegend sind für die Organisation und Analyse von Vektordaten.

Feature Class: Organisationseinheit in ArcGIS (Dataset, Objektthema), das Geoobjekte – Instanzen einer Klasse – mit gleichem Geometrietyp (feature type) enthält.
Ein Geoobjekt wird heute deshalb oft auch als Feature bezeichnet. Dieser Begriff steht aber eigentlich nur für das besondere Merkmal eines Geoobjektes, seine Geometrie, die seine Form und Lage beschreiben.
Der Objektbegriff setzt Identität voraus, die durch eine OID gegeben ist und mittels Geometrie und Sachdaten charakterisiert wird (siehe UML).

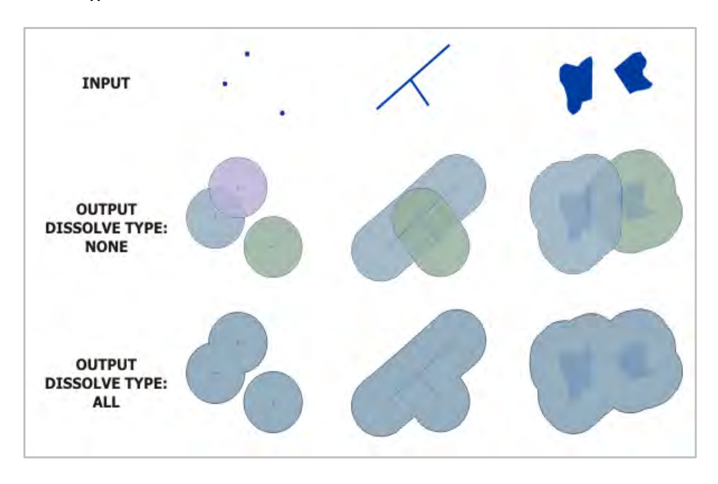
*Geometrische Operatoren erfordern die Eingabe eines oder mehrerer Objekthemen (**in**_feature_class) und erzeugen ein Ergebnisthema (**out**_feature_class)*

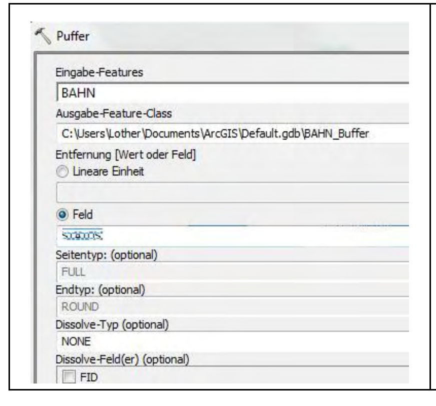
Pufferung von Geoobjekten („Nähe/Distanz von Merkmalen")

ArcToolbox: Analysis Tools > Nachbarschaftsanalyse > Puffer

Puffer sind Zonen (Flächen), die mit einem gegebenen Abstand ein geometrisches Element, wie einen Punkt, eine Linie oder ein Polygon, umschließen. Die Erstellung von Puffern zur Bestimmung der „Nähe" von Merkmalen ist eine häufige Anwendung der GIS-Analyse. Mit Hilfe von Puffern können die räumlichen Einflüsse von Geoobjekten für die Analyse geometrisch dargestellt werden.

Es werden Pufferpolygone mit einem festgelegten Abstand um die Eingabe-Features herum erzeugt. Optional kann eine Zusammenführung (ALL) von Puffern zu einem Polygon erfolgen, wenn sie sich überlappen.





Buffer: Creates buffer polygons to a specified distance around the input features.

```
Buffer <in_features> <out_feature_class> <buffer_distance_or_field> {FULL | LEFT | RIGHT |
  OUTSIDE_ONLY} {ROUND | FLAT} {NONE | ALL | LIST} {dissolve_field;dissolve_field...}
```



- Features will not be buffered if their buffer distance is zero.
- When buffering polygon features, negative distances can be used to create buffers on the inside of the polygon features.

Mehrfachring-Puffer

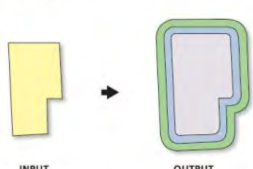
ArcToolbox: Analysis Tools > > Nachbarschaftsanalyse > Mehrfachring-Puffer

Erstellt mehrere Pufferzonen um die Eingabe-Features herum mit den angegebenen Entfernungen.
Diese Puffer können optional mithilfe der Pufferdistanzwerte zusammengeführt werden, um Pufferbereiche zu erstellen, die sich nicht überlappen.

Puffer können in GIS auch verwendet werden, um die Unschärfe von Vektorkonturen analytisch auszuwerten und zu visualisieren.

Multiple Ring Buffer: Creates a new feature class of buffer features using a set of buffer distances.

```
MultipleRingBuffer <input_features> <output_feature_class> <distances;distances...>
  {DEFAULT | CENTIMETERS | DECIMALDEGREES | FEET | INCHES | KILOMETERS | METERS | MILES |
  MILLIMETERS | NAUTICALMILES | POINTS | YARDS} {field_name} {ALL | NONE} {FULL | OUTSIDE_
  ONLY}
```



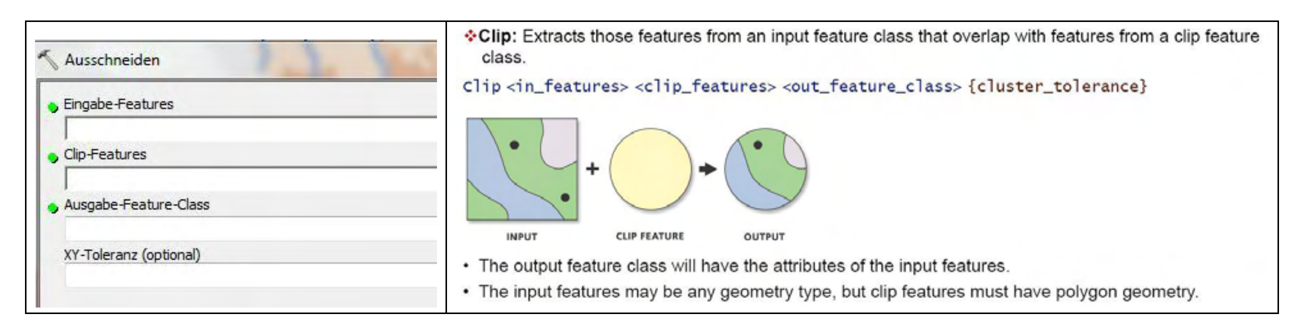
- If the <field_name> is not set, the default name for the field containing the distance value is "distance". The field type is double.

Themen abgrenzen und zusammenführen (Operatoren für die Datenorganisation)

Abgrenzen (Clip)

ArcToolbox: Analysis Tools > Extrahieren > Ausschneiden (clip)

Als Vorbereitung für Analysen ist es oft sinnvoll, die Ausdehnung sehr unterschiedlicher Themen auf das eigentliche Analysegebiet einzuschränken. Dadurch kann das Mitführen überflüssiger Daten im Analyseprozess vermieden werden.



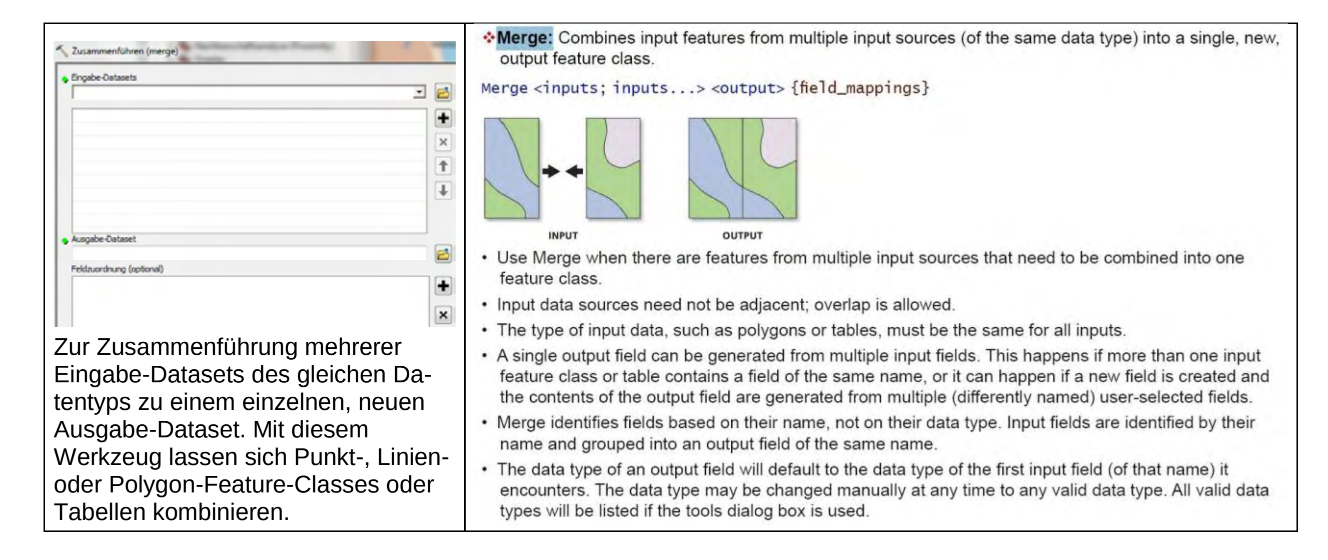
Abgrenzen (Clip): Die Operation „Clip“ verwendet man zum Ausschneiden eines Teils von einem Thema unter Verwendung eines Polygonthemas als „Schablone“.

Dieses Werkzeug kann man verwenden, wenn ein Teil einer Feature-Class mithilfe eines oder mehrerer Features einer anderen Feature-Class (wie mit einer Ausstechform) ausgeschnitten werden soll. Dies ist besonders nützlich beim Erstellen einer neuen Feature-Class (auch als Untersuchungsgebiet oder Interessensbereich bezeichnet), die eine räumliche Teilmenge der Features einer anderen, größeren Feature-Class enthält.

Zusammenführen (Merge)

ArcToolbox: Data Management Tools > Allgemein > Zusammenführen (merge)

Für die Auswertung kann es erforderlich sein, verschiedene Themen zu einem Gesamtthema zusammenzufassen. Mit der Operation Merge kann die räumliche und thematische Vereinigung von (Teil-)Themen gleichen feature types zu einem Gesamtthema erfolgen.



Zur Zusammenführung mehrerer Eingabe-Datasets des gleichen Datentyps zu einem einzelnen, neuen Ausgabe-Dataset. Mit diesem Werkzeug lassen sich Punkt-, Linien- oder Polygon-Feature-Classes oder Tabellen kombinieren.

Zusammenführen (Merge): Die Operation „Merge“ verwendet man, wenn man ein Gesamtthema aus zwei oder mehreren Teilthemen des gleichen Typs erstellen will.

Zerlegen eines Themas

Die umgekehrte Operation kann durch Selektion der Objekte und separieren mit der Funktion: *LY-rM: Daten exportieren* realisiert werden. Das Ergebnis ist ein neues Thema, das nur die ausgewählten Objekte enthält.

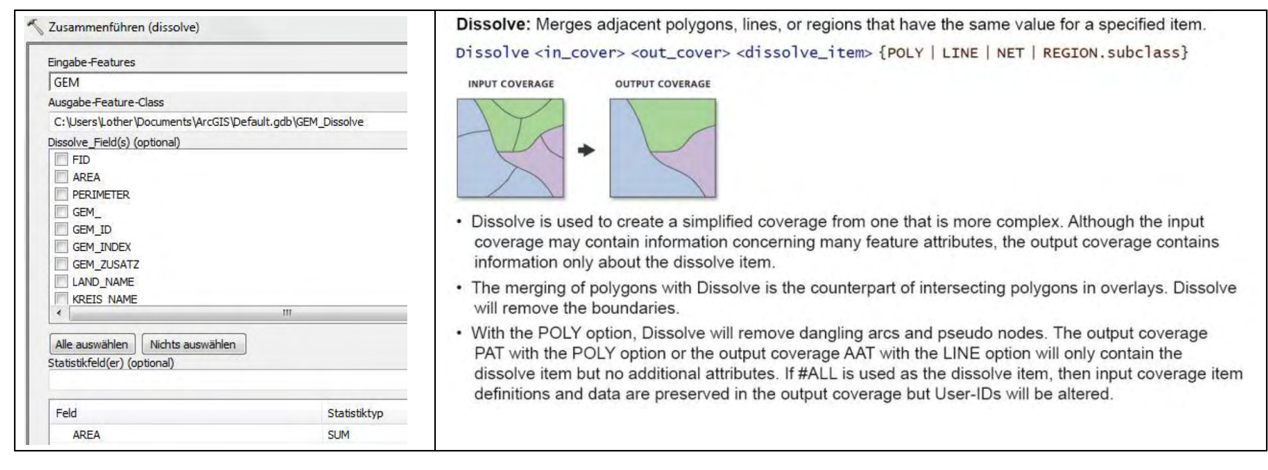
Objekte verschneiden und aggregieren (Grundoperationen für Analysen)

Analyseoperationen ermöglichen es, Geoobjekte in einem oder mehreren Eingabethemen zu verknüpfen. Sie erzeugen ein neues Ausgabethema mit neuen Geoobjekten, die aus der Art der Verknüpfung resultieren. Wichtige Operationen sind die Aggregation von Geoobjekten aufgrund gleicher Attributwerte sowie die Verschneidung im Überlappungsbereich (Intersect) und die Verschneidung mit Überlagerung des Randbereiches des Eingabethemas (Union).

Aggregierung (Dissolve)

ArcToolbox: Analysis Tools > Data Management Tools > Generalisierung > Zusammenführen (dissolve)

Polygone mit denselben Werten im gewählten Aggregationssattribut werden zu komplexen Polygonen vereinigt, ihre gemeinsamen (Innen-)Grenzen werden aufgelöst (*dissolve = auflösen*)



Bei der Aggregierung kann auch festgelegt werden, welche Attribute des Ausgangsthemas zusätzlich aggregiert und in das Ergebnisthema übernommen werden sollen, dies entspricht dem Vorgehen bei der Feldstatistik.

Die "Aggregierung" verwendet man, um Polygone (Flächen) in Abhängigkeit von den Werten eines gewählten Attributs zusammenzufassen. Eingabepolygone mit gleichem Attributwert, die aneinandergrenzen, werden bei der Aggregierung verschmolzen, d.h. die Grenzen zwischen ihnen werden aufgelöst (dissolve). Es können komplexe Polygone entstehen, deren Teilflächen räumlich nicht zusammenhängen, wenn die Option *„Multipart-Geometrie“* gewählt wird.

Eine starke Zersplitterung von Flächenobjekten kann entstehen, wenn mehrere Verschneidungen (intersect) aufeinanderfolgend ausgeführt werden. Die Operation Aggregierung (dissolve) kann man dann einsetzen, um die Teilflächen wieder zu größeren Einheiten zusammenzufassen. Durch die Aggregation kann eine zu starke Zersplitterung von Flächen vermieden werden.

Bei einer Aggregation werden Teile zu komplexen Einheiten (Aggregate), in Abhängigkeit der Werte des gewählten Aggregationsattributs zusammengefasst. In GIS sind verschiedene Arten der Aggregation möglich, die man vom Ergebnis her unterscheiden muss.

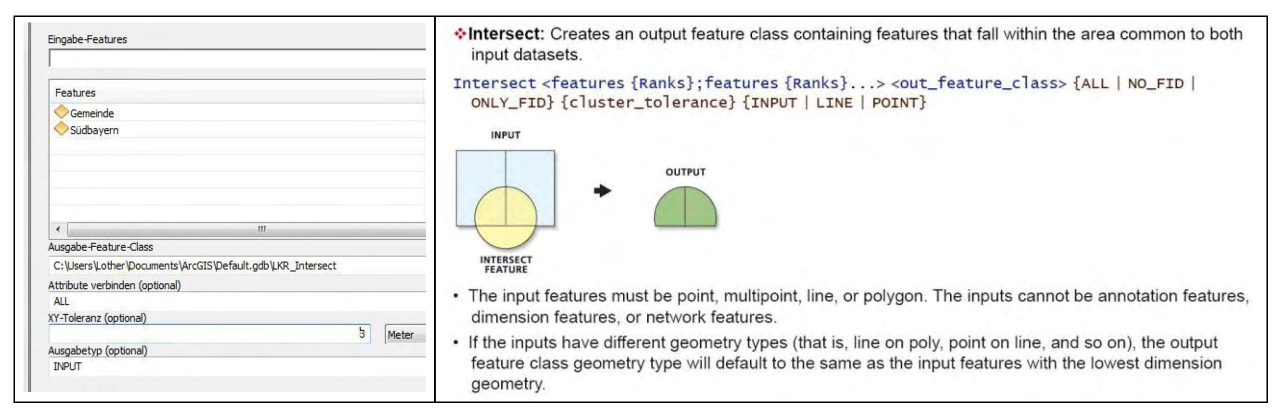
Aggregationstyp	Methode	Ergebnis
Geoobjekte = Geometrie und Attribute	Geoprocessing *„dissolve“* (MP, SP)	neues Thema, mit aggregierten, komplexen Objekten, es werden aggregierte Geometrien gebildet
Attribute	Feldstatistik *„group by“*	neue Tabelle, mit den aggregierten Attributwerten
Grafik = Repräsentation	Symbologie (Style), Legenden vom Typ: *„Einzelwerte“* *„Abgestufte Farbe“*	neue Legende (Style), grafische Aggregation durch Signaturwahl, z.B. Flächenfarbe → erfolgt nur im Bild (View)

Bei Geodaten kann aufgrund der grafischen Darstellung nicht ohne Weiteres unterschieden werden, ob es sich um aggregierte Geoobjekte (Datenstruktur) oder nur um Geoobjekte mit einer klassifizierenden Legende (Grafik) handelt.

Verschneidung (Intersect)

ArcToolbox: Analysis Tools > Overlay > Verschneiden (intersect)

Die Polygone aus den Eingabethemen werden im Überlappungsbereich verschnitten → es werden die Schnittpunkte aller Konturen berechnet und daraus die resultierenden Schnitt-objekte gebildet. Das Ergebnis-Thema enthält neue Geoobjekte, mit den Eigenschaften der Eingabeobjekte in dem Gebiet, das dem Schnittbereich der Eingabethemen entspricht.



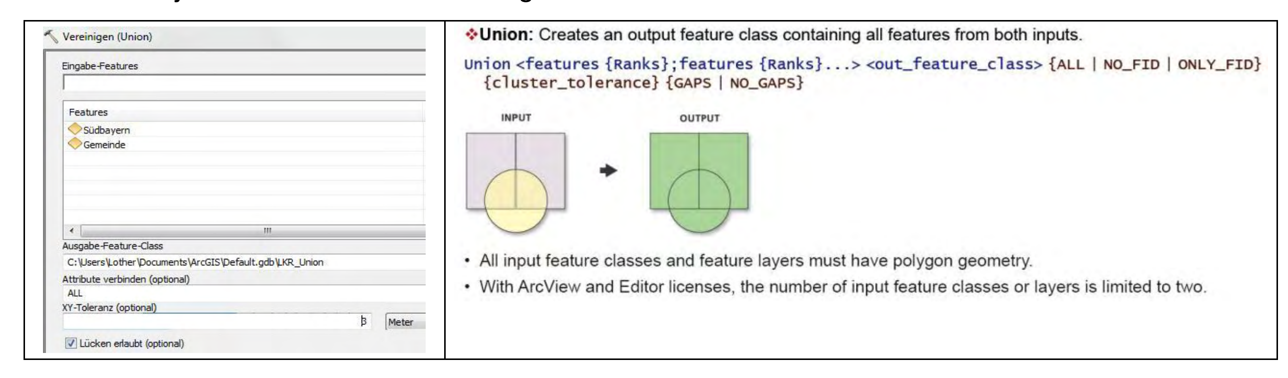
Bei der Intersect-Operation werden Objekte des Eingabethemas, die nicht von Objekten des Intersect-Themas überlagert werden, nicht in das neue Thema aufgenommen. Die Objekte des neuen Themas und die Objekte der Eingabethemen haben denselben Typ. Beim Verschneiden von Objekten des Eingabe-themas mit Polygonen des Überlagerungsthemas werden die Attributtabellen ebenfalls aktualisiert. Die Attributtabelle des Ergebnisthemas enthält die Attribute des Eingabe- und Überlagerungsthemas.

Um ein Verschneidungs-Thema zu erstellen, sind mindestens zwei Themen erforderlich. Das Eingabethema enthält z.B. die Daten der Flächennutzung, das Intersect-Thema (Überlagerungsthema) enthält z.B. die Flurstücksdaten. Als Ergebnis erhält man Teilflächen, die sowohl die Flurstücks- als auch die Flächennutzungsmerkmale besitzen.

Überlagerung (Union)

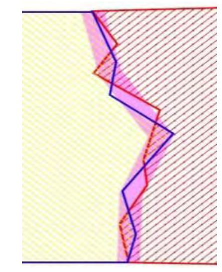
ArcToolbox: Analysis Tools > Overlay > Vereinigen (union)

Die Themen werden im Überlappungsbereich verschnitten, wie beim Intersect. Das Ausgabe-Thema hat jedoch das Gebiet des Eingabe-Themas.



Überlagerung (Union): Die Überlagerung verwendet man, wenn man ein neues Thema erstellen möchten, das alle Objekte und Attribute zweier Polygonthemen enthält. Die Union-Operation kann z.B. zum „Ausschneiden von Inselflächen“ → disjunkte Flächen erzeugen, verwendet werden.

Während bei der Intersect-Operation das Eingabethema mit dem Intersect-Thema verschnitten wird und das Gebiet des Ergebnisthemas dem gemeinsamen Gebiet, das beide Themen abdecken entspricht, wird bei der Union-Operation im Überlappungsbereich der beiden Themen ebenfalls eine Verschneidung durchgeführt, das Ergebnisthema hat jedoch die Ausdehnung des Eingabethemas, d.h. in den Bereichen ohne Überlappung werden die Objekte des Eingabethemas ganz oder angeschnitten übernommen.



Bei den Operationen Intersect und Union ist die „XY-Toleranz“ in Abhängigkeit von der metrischen Genauigkeit der Objektkonturen sinnvoll zu wählen und einzugeben, um Splitterflächen und kleine Lücken im Ergebnis (Slivers / Gaps), die keine signifikanten Objektbereiche darstellen, zu vermeiden.

Intersect von Geodaten

Die Topologische Relation „*intersects*“ und die Geoprocessingoperation (Methode) „*Intersect*“ müssen unterschieden werden!

Topologische Relation „*intersects*“	**Geoprocessing Methode „*Intersect*“**
räumliche Auswahl (Selektion)	Analytische Verschneidung
Fläche **a** Flächen **b**	Beispiel: Thema: Flurstück Thema: Bodenwert Verschneiden → Ergebnisthema: Teilfächen für Einlagewert im Überlagerungs-bereich
Ergebnis **Selektionsmenge** mit den Geoobjekten des Themas b, die sich mit den Geoobjekten des Auswahlthemas a überlappen (overlaps) oder in ihnen enthalten (contains) sind.	Ergebnis **neues Thema** mit den Schnittflächen der Geoobjekte in den Themen a und b. Die neuen Geoobjekte verfügen über alle Attribute der Ausgangsobjekte
Ist die Verwendung einer räumlichen m:n- Relation zur Auswahl (Selektion) von Geoobjekten	Ist Auflösung einer räumlichen m:n- Relation zwischen Geoobjekten → kleinste gemeinsamen Geometrie (kgG)

Ergebnis der „*Intersect*“-Operation

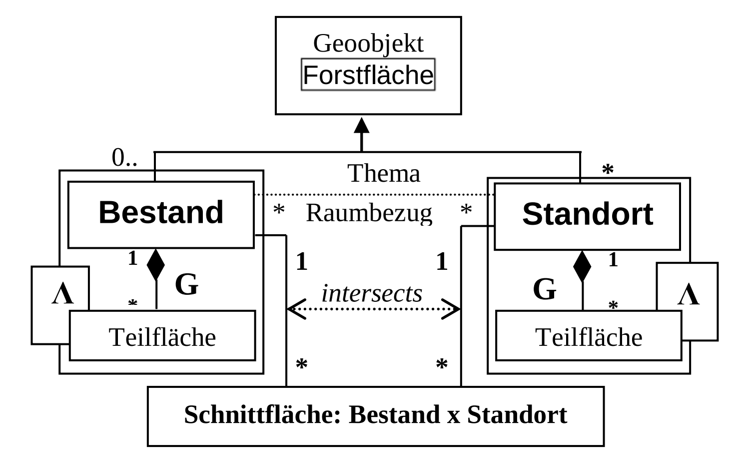
Die geometrischen Beziehungen zwischen verschiedenen, unabhängigen Themen ergeben sich aus dem gemeinsamen Raumbezug der Geometriedaten. Verschiedene Geodaten können aufgrund ihres gemeinsamen Raumbezuges einander überlagert und geometrisch in Beziehung gebracht werden. Mit den oben aufgeführten topologischen Relationen lassen sich dann Abfragen wie: „enthält“, „berührt“, „überlappt“ usw. stellen. Man wählt dann Geobjekte aus, die dieser räumlichen Relation entsprechen. Bei der Verschneidung sucht man dagegen die sogenannte kleinste gemeinsame Geometrie (kgG) der verschnittenen Themen, diese besteht aus solchen Gebieten, die eindeutig allen Eingabethemen zugeordnet werden können.

Ein Bezug der Abfragen auf konkrete Teilflächen, die sich durch Überlappungen von zwei Gruppen von verschiedenen Geoobjekten ergeben, lässt sich nicht herstellen, hierzu ist erst die Verschneidung der Flächen erforderlich.

Für die Modellierung müssen m:n-Relationen auf zwei Relationen 1:n und 1:m zurückgeführt werden, wofür eine neue Tabelle definiert werden muss, die diese Verbindungsrelation aufnimmt.

Für Geoobjekte erfordert die Auflösung einer m:n-Relation die geometrische Verschneidung der Flächen. Die resultierenden Schnittflächen (kgG) haben alle Eigenschaften der Eingabethemen.

Beispiel: Die Beziehung zwischen Waldbeständen und Standorten ist eine m:n-Relation (komplexe Relation).



Verbindung und Verknüpfung von Daten

Für Auswertungen kann es erforderlich sein, dass ein Thema (Geodaten) mit einer Tabelle oder zwei Themen miteinander verbunden werden müssen. Hierfür sind geeignete „Join"-Operationen erforderlich.

Für die Verbindung eines Themas – über seine Attributtabelle – mit einer Tabelle oder zwischen zwei Tabellen stehen die Operationen *„join"* (Verbindung) und *„relate"* (Beziehung) in ArcGIS zur Verfügung. Zwei Themen können zudem über ihre Geometrie mittels einer räumlichen Verbindung *„spatial join"* in Beziehung gesetzt werden.

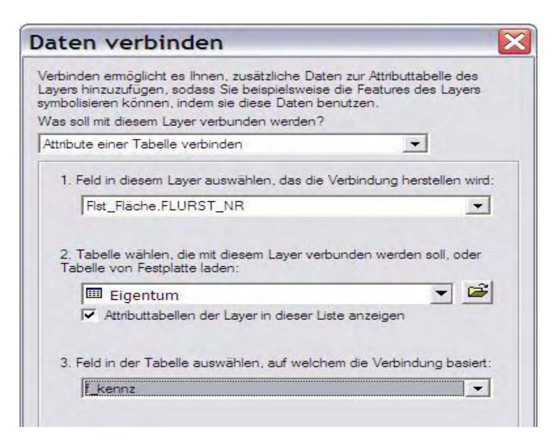
Join, Relate - semantische Verbindung von Tabellen

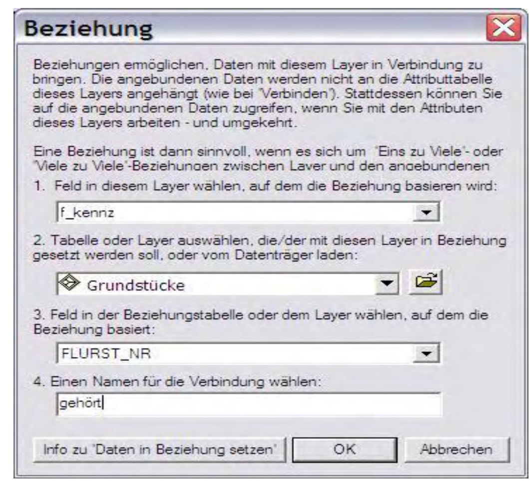
LY-rM → Verbindungen und Beziehungen | TB-Optionen → Verbindungen und Beziehungen

Tabellen können über Attribute (Spalten), die in beiden Tabellen vorkommen und gemeinsame Werte enthalten (gemeinsame Domäne) miteinander in Beziehung gesetzt werden.

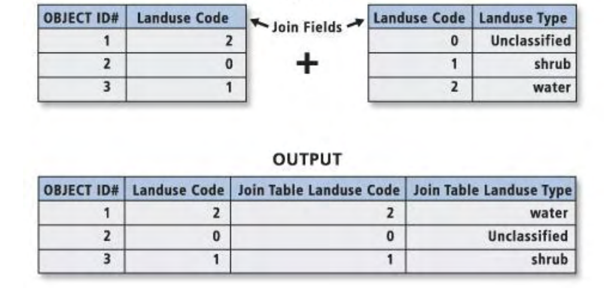
Abhängig vom Beziehungstyp stehen in ArcGIS zwei Funktionen zur Verfügung:

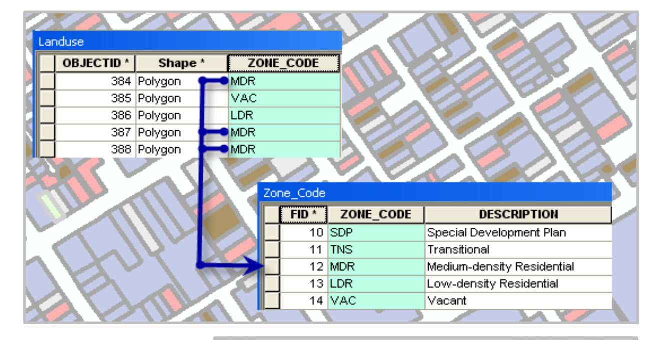
- **join** für **1:1** oder **n:1**, Tabellen verbinden (bidirektionale Beziehung)
- **relate** für **1:n** oder **m:n,** Tabellen in Beziehung setzen (unidirektionale Beziehung)



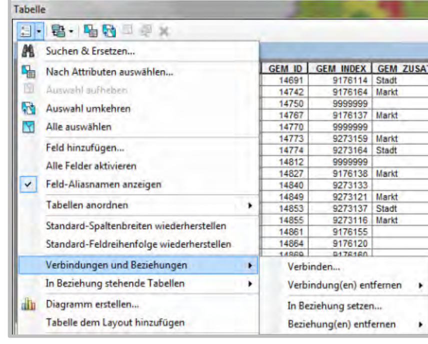


Tabellen kann man verbinden, wenn die Daten der zwei Tabellen eine Eins-zu-eins (1:1) oder Viele-zu-Eins (n:1) Beziehung bilden.
Wenn man zwei Tabellen verbindet, werden im Projekt die Attribute einer Tabelle basierend auf einem gemeinsamen Feld beider Tabellen temporär an die andere Tabelle angefügt.





Man kann Tabellen in Beziehung setzen, wenn ihre Daten eine Eins-zu-viele (1:n) oder Viele-zu-Viele (m:n) Beziehung bilden. Zwei Tabellen in Beziehung zueinander zu setzen heißt, basierend auf einem gemeinsamen Feld entsprechende Tabellenzeilen einander zuzuordnen. Eine Beziehung wird im Projekt jedoch nicht wie eine Verbindung als gemeinsame erweiterte Tabelle (2NF) behandelt, sondern man kann bei Bedarf nur auf die in Beziehung gesetzten Daten zugreifen. Für die bessere Übersicht können Beziehungen mit einem Namen bezeichnet werden.



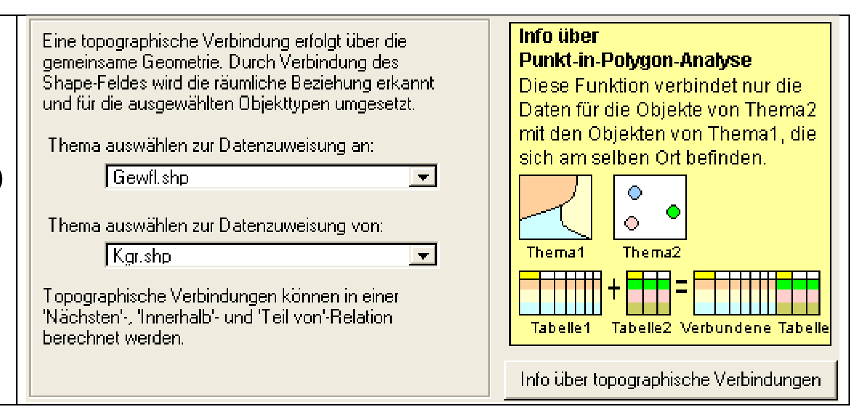
Spatial Join - räumliche Verbindung von Themen

Eine räumliche Verbindung überträgt die Attribute aus einem Thema (Feature-Class) in ein anderes Thema (Feature-Class), basierend auf den räumlichen Beziehungen zwischen den Geoobjekten in den beiden Themen.

Punkt in Polygon-Verbindung

Diese Operation ermöglicht es Geoobjekte in verschiedenen Themen über ihren Raumbezug in Beziehung zu setzen (Relation)

Es kann z.B. eine Beziehung zwischen einem Polygonthema mit Bezugsräumen und einem XY-Thema (Ereignisthema) mit georeferenzierten Fachdaten (Punkte) hergestellt werden.



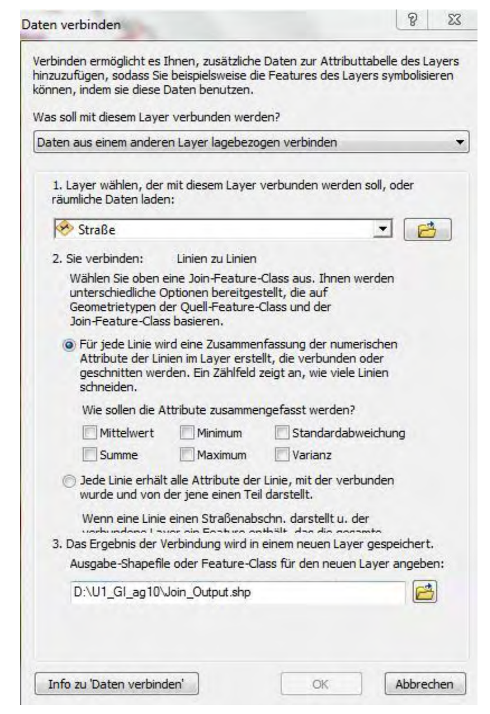
Verbinden von Daten nach Position (räumlich)

LY-rM → Verbindungen und Beziehungen | ArcToolbox: Analysis Tools > Overlay > Räumliche Verbindung

Wenn zwei Themen kein gemeinsames Attributfeld aufweisen, kann man diese mit Hilfe einer räumlichen Verbindung in Beziehung setzen, die Attribute der beiden Themen werden dann basierend auf den Objekt-Positionen der Themen verknüpft. Räumliche Verbindungen unterscheiden sich von Verbindungen über Attribute dadurch, dass sie nicht dynamisch sind und die Ergebnisse in einem neuen Ausgabe-Thema gespeichert werden müssen.

Abhängig vom geometrischen Typ der Daten, handelt es sich bei der Verbindung um eine von drei möglichen Relationen: "am nächsten", "innerhalb" oder "Teil von".

- Zuordnung der einzelnen Features zu den nächstgelegenen Features: Bei dieser Zuordnung können die Attribute des am nächsten gelegenen Features bzw. eine Zusammenfassung der numerischen Attribute der am nächsten gelegenen Features (Minimalwert, Maximalwert usw.) angehängt werden.
- Zuordnung der einzelnen Features zu dem Feature, in dem diese sich befinden: Bei dieser Zuordnung erhalten die Features innerhalb eines Polygons die Attribute der umgebenden Polygonfeatures. Beispiele für eine solche Zuordnung sind ein Punkt innerhalb eines Polygons oder ein Liniensegment, das sich vollständig innerhalb eines anderen Liniensegments befindet (also von diesem überlappt wird).
- Zuordnung der einzelnen Features zu den Features, die von den einzelnen Features überschnitten werden: Wie bei der Zuordnung der oben erwähnten nächstgelegenen Features können die Attribute eines einzelnen überschneidenden Features bzw. eine Zusammenfassung der numerischen Attribute der überschneidenden Features anhängt werden.



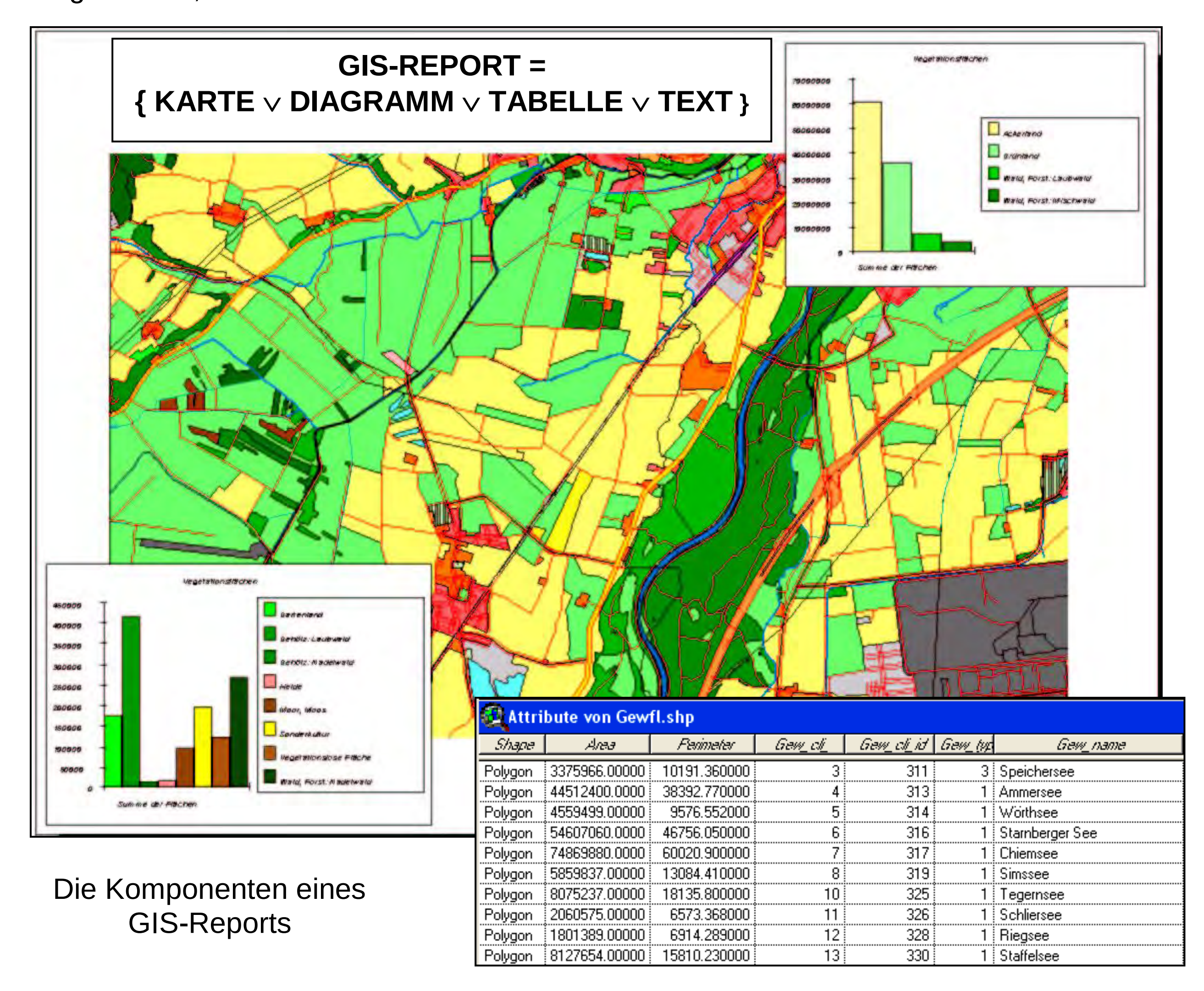
Vereinigungsregeln im Parameter Feldzuordnung von Verbindungs-Features gelten nur für Felder aus den Verbindungs-Features und nur dann, wenn als Verbindungsvorgang JOIN_ONE_TO_ONE verwendet wird. Die Vereinigungsregel wird ignoriert, wenn sie auf Felder aus den Ziel-Features angewendet wird oder wenn als Verbindungsvorgang JOIN_ONE_TO_MANY verwendet wird.

Der Ausgabe-Feature-Class wird stets ein neues Feld mit dem Namen "Join_Count" hinzugefügt. Dieses Feld gibt an, wie viele Verbindungs-Features mit jedem Eingabe-Ziel-Feature verbunden wurden.

Diese Operation benötigt man, wenn man eine raumbezogene Relation (spatial join) verwenden möchte, um Daten aus der Tabelle eines Themas mit der Tabelle eines anderen Themas über die Geometrie der Geoobjekte zu verbinden "Daten nach Ort zuweisen“

Reporterstellung (Präsentation, Visualisierung)

GIS-Produkte können zusammenfassend als Report bezeichnet werden, wobei ein GIS-Report das Ergebnis einer Auswertung (Abfrage oder Analyse) ist und sich aus Karten, Diagrammen, Tabellen und Texten zusammensetzten kann.



Shape	Area	Perimeter	Gew_cl_	Gew_cl_id	Gew_typ	Gew_name
Polygon	3375966.00000	10191.360000	3	311	3	Speichersee
Polygon	44512400.0000	38392.770000	4	313	1	Ammersee
Polygon	4559499.00000	9576.552000	5	314	1	Wörthsee
Polygon	54607060.0000	46756.050000	6	316	1	Starnberger See
Polygon	74869880.0000	60020.900000	7	317	1	Chiemsee
Polygon	5859837.00000	13084.410000	8	319	1	Simssee
Polygon	8075237.00000	18135.800000	10	325	1	Tegernsee
Polygon	2060575.00000	6573.368000	11	326	1	Schliersee
Polygon	1801389.00000	6914.289000	12	328	1	Riegsee
Polygon	8127654.00000	15810.230000	13	330	1	Staffelsee

Die Komponenten eines GIS-Reports

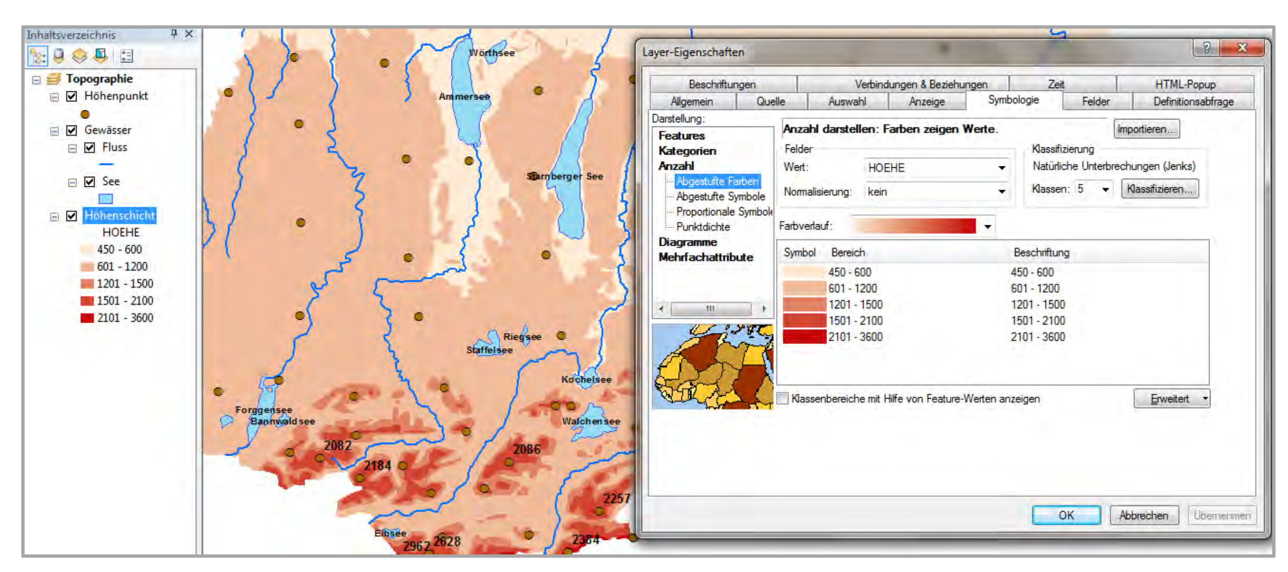
Da mit einem GIS raumbezogene Daten verarbeitet werden, kommt der Karte für die Präsentation von Auswerteergebnissen eine vorrangige Bedeutung zu, häufig ist das Ergebnis einer Auswertung ausschließlich eine Karte (grafische Repräsentation).

Begriffsbestimmung für die Karte nach Salitschew 1967: *"Karten sind verkleinerte, generalisierte, erläuterte und mathematisch bestimmte Abbildungen der Erdoberfläche in der Ebene, welche die Verteilung, den Zustand und die Zusammenhänge der verschiedenen natürlichen und gesellschaftlichen Erscheinungen zeigen, die entsprechend dem Zweck der Karte ausgewählt und charakterisiert werden."*

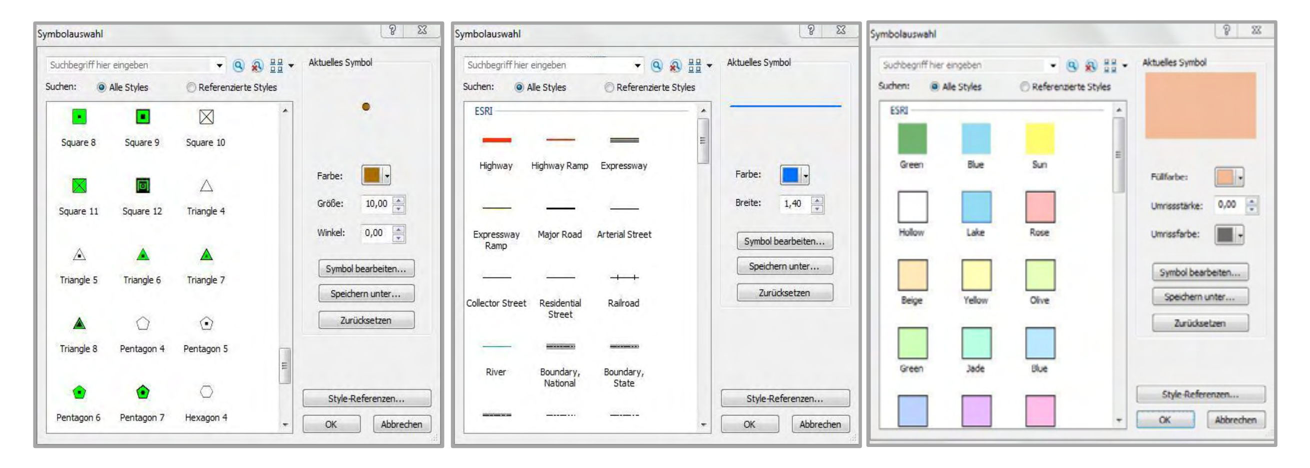
Im GIS entsteht eine Karte als grafische Repräsentation von Geodaten (View), wenn man den Geodaten eine Zeichenanweisung (Style) zuordnet und sie nach dieser Zeichenanweisung grafisch interpretiert. Je nach Zweck und Aufgabe können einem Datenbestand unterschiedliche Styles zugewiesen werden. Je größer der Unterschied zwischen dem Modellmaßstab der Geodaten und dem Ausgabemaßstab der Karte ist, desto umfangreicher sind die anfallenden Arbeitsschritte, um ein zweckorientiertes und lesbares Kartenbild zu erzeugen. Insbesondere können bei stärkeren Verkleinerungen die nötigen Generalisierungen sehr aufwendig werden.

Grafikparameter für die Visualisierung von Geodaten

Die Visualisierung von Geodaten kann als Karte (ArcMap, 2D) oder bei Integration der Höhe als Szene (ArcScene, 3D) erfolgen. Für die Erzeugung einer grafischen Repräsentation von Geodaten wird den Geodaten in GIS-Softwareprodukten mit Hilfe von „Legendeneditoren“ (ESRI: Symbologie) ein Grafik-Style (Legende) zugewiesen.



In ArcGIS werden mit der Layer-Eigenschaft *„Symbologie“* den Features eines Themas (Layer) – je nach Typ: Punkt, Linie, Fläche – Grafikparameter wie Farbe oder Strichbreite zugewiesen sowie die Art der Legende *„Features, Kategorien, Anzahl“* gewählt.



Werkzeuge zum Zuweisen der Grafikparameter

In den meisten GIS-Softwareprodukten hat man drei grundlegende Legendenarten:

- Alle Features eines Themas einheitlich darstellen (Default)
- Features differenziert nach den Werten eines Attributes darstellen, in Abhängigkeit vom Attributtyp
  „Kategorien“ Grafikparameter für Einzelwerte von nominalen, ordinalen oder kardinalen Attributwerten (diskrete Wertemenge)
  „Anzahl“ Grafikparameter für Werteklassen von metrischen Attributwerten (stetige Wertemenge).
- Attributwerte können auch als Texte („Beschriftungen“) visualisiert werden, aus denen sich bei Bedarf unabhängige Textelemente (Annotationen) generieren lassen.

Ab den ArcGIS Versionen 9.2 und 10 wird von ESRI sehr viel Entwicklungsarbeit in die Möglichkeiten der kartographischen Präsentation von Geodaten investiert, um direkt aus Geodaten nicht nur Arbeitskarten (Visualisierungen) sondern auch kartographische Produkte herzuleiten.

Kartographie-Trends (ArcGIS)









Die Vision von ESRI sieht die Zusammenführung aller Workflows und Prozesse für das kartografische Finishing von Kartenprodukten unter einer Oberfläche - der von ArcGIS - vor. Hierzu gehören:

- Eine Arbeitsumgebung von der Erfassung bis zur kartografischen Feinarbeit
- Eine zentrale Geodatenbank
- Multi-Repräsentation von Geodaten für unterschiedliche kartografische Endprodukte
- Intuitive und interaktive Editierwerkzeuge
- Ein System für die Produktion qualitativ hochwertiger Karten

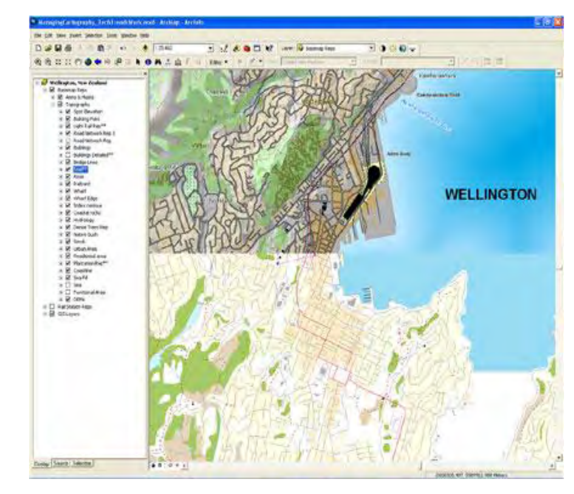
Das Ziel ist es, den Anwendern alle Möglichkeiten an die Hand zu geben, um hochwertige Karten mit klarer Aussage und in präziser Ausführung auf einfache und effiziente Weise zu erzeugen. Damit dies möglich ist, wurde bei ESRI eine ganze Reihe von Werkzeugen zur Automatisierung kartografischer Prozesse entwickelt.

Kartografische (Merfach-)Repräsentationen

Das Konzept der kartografischen (Mehrfach-)Repräsentation ermöglicht es festlegen, wie und wo welches Geoobjekt platziert und auf welche Weise es visualisiert wird.

Mit Repräsentationen kann man:

- aus einem Geodatenbestand viele unterschiedliche Kartenprodukte ableiten
- kartografische Freiheiten auf Geodaten anwenden
- für verschiedene Maßstabsbereiche eigene Kartenprodukte erstellen
- einheitliche Kartenserien definieren



Geoobjekte darstellen

Durch (Mehrfach-)Repräsentation kann man zusätzliche Inhalte mit den Geoobjekten auf der Karte darstellen. Jedem Thema (Feature Class) können mehrere Regelsätze und Ausnahmen von Regeln zugeordnet werden.

Einzelnen Geoobjekten kann man sogar vollkommen freie Darstellungen zuweisen, um auf diese Weise den Inhalt und die Aussage der zweckorientiert optimieren.

In dieser Abbildung sind zwei Ansichten mit Hilfe des "Swipe"-Werkzeuges gemeinsam dargestellt: In der unteren Bildhälfte als einfache grafische Präsentation der Geodaten, in der oberen Bildhälfte eine kartografische Repräsentation der Geodaten.



Generalisierung

Werkzeuge für die automatische bzw. interaktiv unterstützte Generalisierung ermögliche es, die Platzierung, Ausprägung und damit die Visualisierung von Geodaten bei verkleinerten Darstellungen zu verändern. Dies kann aus Platzgründen, aber beim WebMapping auch aus Performance Gründen erfolgen.

Zur Durchführung von Generalisierungsprozessen benötigt man sowohl entsprechende Rechenmodelle als auch Routinen für die Problembehandlung. Beide setzen üblicherweise auf der topologischen Datensicht auf und werden als Prozessschritt vor die Visualisierung gesetzt. Werkzeuge für die gängigsten Generalisierungsaufgaben sind:

- Vereinfachen von Flächen und Linien: Unter Erhalt der charakteristischen Form werden extreme Kurven/Ausbuchtungen herausgerechnet.
- Vereinfachen von Gebäudekonturen: Kleinere Vorsprünge und Gebäudekanten werden fallen gelassen. Die charakteristische Form und Größe der Gebäude bleiben erhalten.
- Zusammenfassen von Flächen: Innerhalb vorgegebener Toleranzgrenzen werden Flächen zusammengefasst.
- Doppellinien zur Mittellinie zusammenfassen: Bei Straßen oder Gewässern kann mit diesem Werkzeug aus beidseitig erfassten Objektgrenzen die Mittellinie errechnet werden.

Übungen zum Modul Geoinformatik

Übung	Thema	Lerninhalte
0 **(2 h)** **Produkt**	**Einführung in ArcGIS** **Datenorganisation, wichtige Funktionen und Customizing** *„Einweisung in die Bedienung von ArcGIS/Base (10.x)“*	ArcMap starten, neues Map Document anlegen • Ordner mit Geodaten im ArcCatalog verbinden • Eigenschaften des Map Documents festlegen • Geodaten hinzufügen • Eigenschaften Data Frame, Layer einstellen • Organisation der Daten (Projekt | Data Frame | Layer | Group Layer) • Bildmanipulationen (Sichtbarkeit, Reihenfolge, Zoom) • Grafikmerkmale, Symbologie (Varianten) • Selektion attributiv, geometrisch
1 **(6 h)**	**Vektordaten** **Geoobjekte, Objektthemen** *weitere Einführung in die Grundfunktionen von ArcGIS (2h)* *„Kennen lernen der Organisation von Vektordaten, Selektion, Erstellung von Karten (LayOut)“* ***DS_Vektor***	Vektorthema (ESRIS/Shape) • Zusammenhang Geometrie- und Sachdaten • Geometrische Selektion • Auswahl nach Attributen, attributive Selektion • Attribute in Tabelle einfügen • Attributwerte editieren und berechnen • Verbindung von Tabellen (Join) • Text-Darstellung von Attributwerten • Texte anpassen (Inhalt, Position, Format) • Layout für Karte erstellen, gestalten
2 **(6 h)**	**Rasterdaten** **Bildthemen, Bilder** *„Bildthemen verarbeiten, Unterschiede zwischen Vektor- und Rasterdaten kennen lernen, Hyperlinks für Multimedia-Attribute“* ***DS_Raster (Bild)***	Bildthema, georeferenzierte Rasterdaten • Vergleich Vektor- mit Rasterdaten • Hybride Datenverarbeitung • Maßstabsabhängige Darstellung von Themen • Importfilter für Raster- und Vektordaten • Vektorthema (AutoCad/DXF) • Objektbearbeitung, erzeugen und ändern • OnScreen Digitalisierung • Hyperlinks für „Multimedia-Attribute“
3 **(6 h)**	**Höhenintegration, DGM** *„3D-Präsentation von Geodaten sowie die Ableitung von höhen-relevanten 2D-Themen unter Verwendung von amtlichen Geobasisdaten (DGM5, DOP, ATKIS/DLM25) aufzeigen.“* ***DS_TIN/GRID***	DGM-Daten (DGM5) • ASCII-Daten in MS-Access einlesen • Ereignisthema aus Access-Tabelle aufbauen • ArcScene starten, neues Scene-Document anlegen • TIN erzeugen • Darstellungsvarianten Höhe, Neigung, Exposition • 3D-Szene erzeugen, Höhenreferenz • 2D-Themen überlagern (DOP oder ATKIS) • 2D-Themen rastern → GRID • DGM-Produkte für 2D-Analyse
4 **(6 h)**	**Geoprocessing** **für Flächenobjekte (Forst)** *„Auflösung der m:n-Relation zwischen Geoobjekten sowie die Aggregation von Geoobjekten; Unterschiede zwischen Primär- und Sekundärmodell (Grafik) aufzeigen“*	Auswertung / Analyse • Umfangreichere Geodaten verwenden • Tabellendarstellung anpassen, Werte bearbeiten • View-Legenden importieren • Aggregation von Geoobjekten • Komplexe Geometrien • Verschneidung von Geoobjekten (intersect) kgG (kleinste gemeinsame Geometrie) • Hintergrund, geeigneter Raumbezug für Geofachdaten

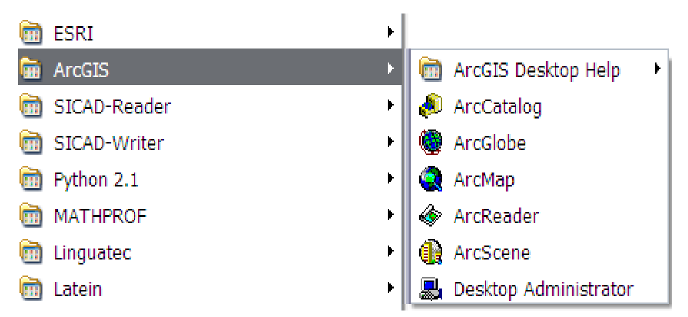
(Zusammenfassung zum Übungsstoff Geoprocessing im Skriptum GI, Kap. 7)

Übungen zum Modul Geoinformationssysteme

Übung	Thema	Lerninhalte
5 **(6 h)**	**Georeferenzierung** **von Rasterdaten** *„Projektion von Vektordaten, Einpassung und Georeferenzierung von Rasterdaten“*	Rasterdaten einpassen ins Weltkoordinatensystem • Worldfile-Standard für Rasterdaten • Geographische Koordinaten für Themen • Projektionen (Kartographische Abbildungen) und Transformation von Koordinaten • Gescannte Rasterdaten georeferenzierten
6 **(6 h)**	**Geobasisdaten, ATKIS** **Auswertebeispiele für unterschiedliche Objekttypen** *„Am Beispiel der Geobasisdaten ATKIS werden Berechnungs- und Aggregationsfunktionen für Tabellen gezeigt, als Grundlage für Diagramme“*	ATKIS-Objektstruktur, Abbildung in Shape-Format • Berechnung von metrischen Größen (Fläche, Länge) • Statistik für numerische Attribute • Feldstatistik zur Aggregation nach Attributwerten • Diagramme erstellen • Objekt-Themen splitten / zusammenfügen (merge) • Vergleich von Vektor- mit Rasterdaten (ATKIS/DTK) • Farbwahl für binäre Rasterbilder • Layout mit Diagrammen
6a **(4 h)**	**Pufferfunktion** *Beispiel: Naturschutz (FFH), ATKIS* *„Pufferfunktion zur geometrischen Repräsentation von entfernungs-abhängigen Einflüssen von Geoobjekten“*	ATKIS-Projekt, Pufferung von Geoobjekten (Gesamtpuffer) • Projekt, Arbeitsvorbereitung (selbständig) • Pufferparameter • Reporterstellung, Karte, Diagramme, Texte • Objekterfassung, Flächenberechnung • Reporterstellung, Kartenexport (selbständig)
7 **(4 h)**	**Räumliche Abfragen** *Beispiel: Geomarketing* *„semantisches Modell UML“* *„Pufferfunktion für entfernungs-abhängige Einflüsse von Geoobjekten und räumliche Abfragen“*	Kommunal-Projekt, Pufferung von Geoobjekten (Einzelpuffer) • Datenrecherche, semantisches Modell (UML) • Logisches Modell, Datenstrukturen • Pufferung, entfernungsabhängige Einflüsse von Geoobjekten • Verknüpfung von Tabellen mit Themen (join, relate) • Geocodierung von Sachdaten • Räumliche Abfragen mittels topologischer Relationen • Kombinierte Abfragen, räumlich und attributiv • Reporterstellung, Karte, Diagramme, Texte
S **(4+2 h)**	**GIS-Software** *Internet-Recherche* *„Wichtige GIS-Produkte/-Hersteller“* *Ergebnis: je Gruppe ein Handout und ein Kurzvortrag.*	Web-Seiten der wichtigsten GIS-Softwarefirmen recherchieren: • Hersteller-Sitz • Niederlassung in Deutschland / Bayern / Raum München • Produkte, Typ der Produkte • Einsatzschwerpunkte • Verbreitung, Marktanteil • Datenhaltungskonzept, Datenschnittstellen • Internetkomponente (WebGIS)
R **(4+2 h)**	**openGIS** *Internet-Recherche* *„Status der GDI aktuell recherchieren“* *Ergebnis: je Gruppe ein Handout und ein Kurzvortrag.*	Web-Seiten mit wichtigsten GDI-Themen recherchieren: • Internationale Standards und Normen (OGC/ISO/W3C) • Geobasisdaten online, Beispiel BayernAtlas • GeoPortale Bund und Bayern • Geodateninfrastrukturen (INSPIRE, GDI-DE, GDI-BY)

ArcView – Desktop GIS for Mapping, Data Integration, and Analysis

ArcGIS Desktop – ArcView – Programmkomponenten



ArcCatalog = ein spezieller Dateiexplorer für Geodaten, mit dem Geodaten verwaltet werden können und alle anderen Dateien im Dateiverwaltungssystem „ausgeblendet" werden können, man sieht nur noch die Geodaten.

Er unterstützt die speziellen Eigenschaften von Geodaten, d.h. Geodaten können

- *grafisch dargestellt werden*
- *es können Metadaten hinterlegt werden*
- ***das Koordinatensystem definiert werden***
- *Verbindungen zu Datenbanken und WMS.*

Der Zugriff über Netz auf die Daten dauert teilweise sehr lange. → Abhilfe: Für die Durchführung der Übungen empfiehl es sich die Ordner mit den Ausgangsdaten auf eine CD zu brennen oder auf einen Speicherstick zu kopieren, um sie während der Übung lokal am Rechner benutzen zu können.

Bei der Migration der Übungen von AV3 → AG9 die Übungsvorlagen so erstellen, dass die Aufgabe getrennt wird von den anschließenden Hinweisen zur Ausführung der Übungen, die nur in analoger Form an die Übungsteilnehmer ausgegben werden → keine digitale Version mehr bereitstellen.

Die Tastenkombinationen „Strg+C" (kopieren) und „STRG+V" (einfügen) sind sehr hilfreich. Die Einfügefunktion kann immer verwendet werden, auch wenn die Zwischenablage mit einer anderen Funktion gespeichert wurde → ArcScene: „in Zwischenablage Speichern", dann „STRG+V" zum Einfügen möglich.

ArcGIS Desktop – Leistungsstufen *(ab v10.1: Base – Standard – Extended)*

Für die Hausübung wird den Studenten die ArcGIS Desktop 10 Education DVD zur Verfügung gestellt

ArcGIS Desktop besteht aus einer Reihe von Software Produkten, mit denen Sie raumbezogene Daten erzeugen, vielfältig nutzen und ausgeben können. ArcGIS Desktop Produkte sind:

- ***ArcInfo (→ Extended)*** *ist der de-facto Standard für GIS Profis. In dieser Ausbaustufe sind umfangreiche Werkzeuge für die Automatisierung von Prozessen, die räumliche Modellierung, Analyse und kartografische Aufbereitung von Daten enthalten. ArcInfo unterstützt modellbasierte Datenerfassung im Mehrbenutzerbetrieb genauso wie die volle Implementierung und Administration der Geodatabase*
- ***ArcEditor (→ Standard)*** *fasst alle für anspruchsvolle Erfassungsaufgaben benötigten Funktionen zusammen. Dies stellt die Datenintegrität sicher und erlaubt Ihnen die Abbildung entsprechender Workflows. Nutzen Sie ArcEditor auch, um Ihre Daten zu verwalten, zu analysieren und für die Ausgabe aufzubereiten.*
- ***ArcView (→ Base)*** *dient vornehmlich der Visualisierung und explorativen Analyse Ihrer Daten. Einfache Erfassung und Ausgaben von Karten und Grafiken sind mit ArcView möglich wie die Nutzung der analytischen Funktionen der verschiedenen optionalen Erweiterungen. ArcView ist das mit deutlichem Abstand verbreitetste und bekannteste GIS Produkt weltweit.*
- ***ArcReader*** *rundet als kostenfreier Viewer die Palette der ArcGIS Desktop Produkte ab. Mit ArcInfo, ArcEditor oder ArcView erstellte Karten werden mit ArcReader für weite Nutzerkreise zugänglich gemacht. Funktional sind Kartenabfragen, die Visualisierung und die Ausgabe von Karten möglich.*
- ***ArcGIS Desktop Erweiterungen*** *fassen jeweils zusätzliche Funktionsblöcke wie etwa Raster-Analyse, 3D-Analyse und Visualisierung oder Geostatistik zusammen.*

Zu den ArcGIS Desktop Produkten gehören jeweils die beiden Anwendungen ArcMap mit ArcToolbox und ArcCatalog. Damit sorgen einheitliche Benutzeroberflächen und eine stets gleichartige Entwicklungs-umgebung für den leichten Austausch von Dokumenten und Entwicklungen. Karten, Daten, Symbolbiblio-theken, Geoverarbeitungsmodelle (aus dem ModelBuilder), eigene Werkzeuge, Schnittstellen, Berichte, Diagramme und Metadaten können standardisiert ausgetauscht werden.